

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT: HUSE, WILLIAM D.
- (ii) TITLE OF INVENTION: SURFACE EXPRESSION LIBRARIES OF HETEROMERIC RECEPTORS
- (iii) NUMBER OF SEQUENCES: 75
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: PRETTY, SCHROEDER, BRUEGGEMANN & CLARK
  - (B) STREET: 444 SO. FLOWER STREET, SUITE 200
  - (C) CITY: LOS ANGELES
  - (D) STATE: CALIFORNIA
  - (E) COUNTRY: UNITED STATES
  - (F) ZIP: 90071
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER:
  - (B) FILING DATE:
  - (C) CLASSIFICATION:
- (vii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: CAMPBELL, CATHRYN A.
  - (B) REGISTRATION NUMBER: 31,815
  - (C) REFERENCE/DOCKET NUMBER: P31 8882
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: 619-535-9001
  - (B) TELEFAX: 619-535-8949

## (2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 7445 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: both
  - (D) TOPOLOGY: circular

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AATGCTACTA CTATTAGTAG AATTGATGCC ACCTTTTCAG CTCGGGCCCC AAATGAAAAT	60
ATAGCTAAAC AGGTTATTGA CCATTTCGGA AATGTATCTA ATGGTCAAAC TAAATCTACT	120
CGTTCGCAGA ATTGGGAATC AACTGTTACA TGGAAATGAAA CTTCAGACA CCGTACTTTA	180
GTTCATATT TAAAACATGT TGAGCTACAG CACCAGATTC AGCAATTAAG CTCTAAGCCA	240
TCTGCAAAAA TGACCTCTTA TCAAAAGGAG CAATTAAAGG TACTCTCTAA TCCTGACCTG	300
TTGGAGTTTG CTTCCGGTCT GGTTCGCTTT GAAGTCCGAA TTAAAACGCG ATATTTGAAG	360
TCTTTCGGGC TTCCTCTTAA TCTTTTGTAT GCAATCCGCT TTGCTTCTGA CTATAATAGT	420

CAGGGTAAAG ACCTGATTTT TGATTTATGG TCATTCTCGT TTTCTGAACT GTTTAAAGCA	480
TTTGAGGGGG ATTCAATGAA TATTTATGAC GATTCCGCAG TATTGGACGC TATCCAGTCT	540
AAACATTTTA CTATTACCCC CTCTGGCAA ACTTCTTTTG CAAAAGCCTC TCGCTATTTT	600
GGTTTTATC GTCGTCTGGT AAACGAGGGT TATGATAGTG TTGCTCTTAC TATGCCTCGT	660
AATTCCTTTT GCGGTTATGT ATCTGCATTA GTTGAATGTG GTATTCCTAA ATCTCAACTG	720
ATGAATCTTT CTACCTGTAA TAATGTTGTT CCGTTAGTTC GTTTTATTAA CGTAGATTTT	780
TCTTCCCAAC GTCCTGACTG GTATAATGAG CCAGTTCCTA AAATCGCATA AGGTAATTCA	840
CAATGATTAA AGTTGAAATT AAACCATCTC AAGCCCAATT TACTACTCGT TCTGGTGTTC	900
CTCGTCAGGG CAAGCCTTAT TCACTGAATG AGCAGCTTTG TTACGTTGAT TTGGGTAATG	960
AATATCCGGT TCTTGTCAAG ATTACTCTTG ATGAAGGTCA GCCAGCCTAT GCGCCTGGTC	1020
TGTACACCGT TCATCTGTCC TCTTTCAAAG TTGGTCAGTT CGGTTCCTT ATGATTGACC	1080
GTCTGCGCCT CGTTCGGGCT AAGTAACATG GAGCAGGTCC CGGATTTCTGA CACAATTTAT	1140
CAGGCGATGA TACAAATCTC CGTTGTACTT TGTTCGCGC TTGGTATAAT CGCTGGGGGT	1200
CAAAGATGAG TGTTTTAGTG TATTCCTTCG CCTCTTCCTT TTTAGGTTGG TGCCTTCGTA	1260
GTGGCATTAC GTATTTTACC CGTTTAATGG AACTTCCTC ATGAAAAAGT CTTTAGTCCT	1320
CAAAGCCTCT GTAGCCGTTG CTACCCTCGT TCCGATGCTG TCTTTCGCTG CTGAGGGTGA	1380
CGATCCCGCA AAAGCGGCCT TTAAGTCCCT GCAAGCCTCA GCGACCGAAT ATATCGGTGA	1440
TGCGTGGGGC ATGGTTGTTG TCATTGTCGG CGCAACTATC GGTATCAAGC TGTTTAAGAA	1500
ATTCACCTCG AAAGCAAGCT GATAAACCGA TACAATTAAA GGCTCCTTTT GGAGCCTTTT	1560
TTTTTGAGA TTTTCAACGT GAAAAAATTA TTATTCGCAA TTCCTTTAGT TGTTCCCTTC	1620
TATTCTCACT CCGCTGAAAC TGTGAAAGT TGTTTAGCAA AACCCCATAC AGAAAATTCA	1680
TTTACTAAGC TCTGGAAAGA CGACAAACT TTAGATCGTT ACGCTAACTA TGAGGGTTGT	1740
CTGTGGAATG CTACAGGCGT TGTAGTTTGT ACTGGTGACG AAATCAGTG TTACGGTACA	1800
TGGGTTCTTA TTGGGCTTGC TATCCCTGAA AATGAGGGTG GTGGCTCTGA GGGTGGCGGT	1860
TCTGAGGGTG GCGGTTCTGA GGGTGGCGGT ACTAAACCTC CTGAGTACGG TGATACACCT	1920
ATTCGGGGCT ATACTTATAT CAACCCTCTC GACGGCACTT ATCCGCCTGG TACTGAGCAA	1980
AACCCCGCTA ATCCTAATCC TTCTCTTGAG GAGTCTCAGC CTCTTAATAC TTTGATGTTT	2040
CAGAATAATA GGTTCGAAA TAGGCAGGGG GCATTAACTG TTTATACGGG CACTGTTACT	2100
CAAGGCACTG ACCCCGTAA AACTTATTAC CAGTACACTC CTGTATCATC AAAAGCCATG	2160
TATGACGCTT ACTGGAACGG TAAATTCAGA GACTGCGCTT TCCATTCTGG CTTTAATGAA	2220
GATCCATTCT TTTGTGAATA TCAAGGCCAA TCGTCTGACC TGCCTCAACC TCCTGTCAAT	2280
GCTGGCGGCG GCTCTGGTGG TGGTCTCGT GCGGCTCTG AGGGTGGTGG CTCTGAGGGT	2340
GGCGGTTCTG AGGGTGGCGG CTCTGAGGGA GCGGTTCCG GTGGTGGCTC TGGTCCGGT	2400
GATTTTGATT ATGAAAAGAT GGCAAACGCT AATAAGGGGG CTATGACCGA AAATGCCGAT	2460

GAAAACGGCG TACAGTCTGA CGCTAAAGGC AAACCTTGATT CTGTCGCTAC TGATTACGGT	2520
GCTGCTATCG AIGGTTTCAT TGGTGACGTT TCCGGCCTTG CTAATGGTAA TGGTGCTACT	2580
GGTGATTTTG CTGGCTCTAA TTCCCAAATG GCTCAAGTCG GTGACGGTGA TAATTCACCT	2640
TTAATGAATA ATTTCCGTCA ATATTTACCT TCCCTCCCTC AATCGGTTGA ATGTCGCCCT	2700
TTTGTCTTTA GCGCTGGTAA ACCATATGAA TTTTCTATTG ATTGTGACAA AATAAACTTA	2760
TTCCGTGGTG TCTTTGCGTT TCTTTTATAT GTTGCCACCT TTATGTATGT ATTTTCTACG	2820
TTTGCTAACA TACTGCGTAA TAAGGAGTCT TAATCATGCC AGTTCTTTTG GGTATTCCGT	2880
TATTATTGCG TTTCCTCGGT TTCCTTCTGG TAACTTTGTT CGGCTATCTG CTTACTTTTC	2940
TTAAAAAGGG CTTGCGTAAG ATAGCTATTG CTATTTCAAT GTTTCTTGCT CTTATTATTG	3000
GGCTTAACTC AATTCTTG TGTTATCTCT CTGATATTAG CGCTCAATTA CCCTCTGACT	3060
TTGTTACAGG TGTTACAGTA ATTCTCCCGT CTAATGCGCT TCCCTGTTTT TAIGTTATTC	3120
TCTCTGTAAA GGCTGCTATT TTCATTTTTG ACGTTAAACA AAAAATCGTT TCTTATTTGG	3180
ATTGGGATAA ATAATATGGC TGTTTATTTT GTAAGTGGCA AATTAGGCTC TGGAAAGACG	3240
CTCGTTAGCG TTGGTAAGAT TCAGGATAAA ATTGTAGCTG GGTGCAAAAT AGCAACTAAT	3300
CTTGATTTAA GGCTTCAAAA CCTCCCGCAA GTCGGGAGGT TCGCTAAAAC GCCTCGCGTT	3360
CTTAGAATAC CGGATAAGCC TTCTATATCT GATTIGCTTG CTATTGGGCG CGGTAATGAT	3420
TCCTACGATG AAAATAAAAA CGGCTTGCTT GTTCTCGATG AGTGCGGTAC TTGGTTTAAT	3480
ACCCGTCTT GGAATGATAA GGAAAGACAG CCGATTATTG ATTGGTTTCT ACATGCTCGT	3540
AAATTAGGAT GGGATATTAT TTTTCTTGTT CAGGACTTAT CTATTGTTGA TAAACAGGCG	3600
CGTTCTGCAT TAGCTGAACA TGTTGTTTAT TGTCGTCGTC TGGACAGAAT TACTTTACCT	3660
TTTGTCGGTA CTTTATATTG TCTTATTACT GGCTCGAAAA TGCCTCTGCC TAAATTACAT	3720
GTGCGGTTG TTAAATATGG CGATTCTCAA TTAAGCCCTA CTGTTGAGCG TTGGCTTTAT	3780
ACTGGTAAGA' AATTGTATAA CGCATATGAT ACTAAACAGG CTTTTTCTAG TAATTATGAT	3840
TCCGGTGTIT AATCTTATTT AAGGCCTTAT TTATCACACG GTCGGTATTT CAAACCATTA	3900
AATTTAGGTC AGAAGATGAA GCTTACTAAA ATATATTTGA AAAAGTTTTT ACGCGTCTT	3960
TGTCTTGCGA TTGGATTTGC ATCAGCATTT ACATATAGTT ATATAACCCA ACCTAAGCCG	4020
GAGGTTAAAA AGGTAGTCTC TCAGACCTAT GATTTTGATA AATTCATAT TGAATCTTCT	4080
CAGCGTCTTA ATCTAAGCTA TCGCTATGTT TTCAAGGATT CTAAGGGAAA ATTAATTAAT	4140
AGCGACGATT TACAGAAGCA AGGTIATTCA CTCACATATA TTGATTTATG TACTGTTTCC	4200
ATTAATAAAG GTAATTCAAA TGAAATTGTT AAATGTAATT AATTTTGTTT TCTTGATGTT	4260
TGTTTCATCA TCTTCTTTTG CTCAGGTAAT TGAAATGAAT AATTCGCCCT TCGCGGATTT	4320
TGTAACCTTG TATTCAAAGC AATCAGGCGA ATCCGTATT GTTTCTCCCG ATGTAAAAGG	4380
TACTGTTACT GTATATTCAT CTGACGTAA ACCTGAAAAT CTACGCAATT TCTTTAFTTC	4440
TGTTTTACGT GCTAATAATT TTGATATGGT TGTTTCAATT CCTTCCATAA TTCAGAAGTA	4500

TAATCCAAAC AATCAGGATT ATATTGATGA ATTGCCATCA TCTGATAATC AGGAATATGA	4560
TGATAAATCC GCTCCTTCTG GTGGTTTCTT TGTTCCGCAA AATGATAATG TTAATCAAAC	4620
TTTTAAATTT AATAACGTTT GGGCAAAGGA TTTAATACGA GTTGTGCAAT TGTTCGTAAA	4680
GTCTAATACT TCTAAATCCT CAAATGTATT ATCTATTGAC GGCTCTAATC TATTAGTTGT	4740
TAGTGCACCT AAAGATATTT TAGATAACCT TCCTCAATTC CTTTCTACTG TTGATTGGCC	4800
AACTGACCAG ATATTGATTG AGGGTTTGAT ATTTGAGGTT CAGCAAGGTG ATGCTTTAGA	4860
TTTTTCATTT GCTGCTGGCT CTCAGCGTGG CACTGTTGCA GGGGGTGTTA ATACTGACCG	4920
CCTCACCTCT GTTTTATCTT CTGCTGGTGG TTCGTTCCGT ATTTTAAATG GCGATGTTTT	4980
AGGGCTATCA GTTCGCGCAT TAAAGACTAA TAGCCATTCA AAAATATTGT CTGTGCCACG	5040
TATTCTTACG CTTTCAGGTC AGAAGGGTTC TATCTCTGTT GGCCAGAATG TCCCTTTTAT	5100
TACTGGTCGT GTGACTGGTG AATCTGCCAA TGTAATAAT CCATTTGAGA CGATTGAGCG	5160
TCAAAATGTA GGTATTTCCA TGAGCGTTTT TCCTGTTGCA ATGGCTGGCG GTAATATTGT	5220
TCTGGATATT ACCAGCAAGG CCGATAGTTT GAGTTCTTCT ACTCAGGCAA GTGATGTTAT	5280
TACTAATCAA AGAAGTATTG CTACAACGGT TAATTTGCGT GATGGACAGA CTCTTTTACT	5340
CGGTGGCCTC ACTGATTATA AAAACACTTC TCAAGATTCT GGGGTACCGT TCCTGTCTAA	5400
AATCCCTTTA ATCGGCCTCC TGTTTAGCTC CCGCTCTGAT TCCAACGAGG AAAGCACGTT	5460
ATACGTGCTC GTCAAAGCAA CCATAGTACG CGCCCTGTAG CGGCGCATTG AGCGCGGCGG	5520
GTGTGGTGGT TACGCGCAGC GTGACCGCTA CACTTGCCAG CGCCCTAGCG CCCGCTCCTT	5580
TCGCTTCTT CCCTTCCTTT CTCGCCACGT TCGCCGGCTT TCCCCGTCAA GCTCTAAATC	5640
GGGGGCTCCC TTTAGGGTTC CGATTTAGTG CTTTACGGCA CCTCGACCCC AAAAACTTG	5700
ATTGGGGTGA TGTTTCACGT AGTGGGGCAT CGCCCTGATA GACGGTTTTT CGCCCTTTGA	5760
CGTTGGAGTC CACGTTCCTT AATAGTGGAC TCTTGTTCCT AACTGGAACA AACTCAACC	5820
CTATCTCGGGCTATTCTTTT GATTTATAAG GGATTTTGGC GATTTGGGAA CCACCATCAA	5880
ACAGGATTTT CGCCTGCTGG GGCAAACCAG CGTGGACCGC TTGCTGCAAC TCTCTCAGGG	5940
CCAGGCGGTG AAGGGCAATC AGCTGTTGCC CGTCTCGCTG GTGAAAAGAA AAACCACCTT	6000
GGGGCCCAAT ACGCAAACCG CCTCTCCCCG CGCGTTGGCC GATTCATTAA TGCAGCTGGC	6060
ACGACAGGTT TCCCGACTGG AAAGCGGGCA GTGAGCGCAA CGCAATTAAT GTGAGTTAGC	6120
TCACTCATTG GGCACCCACG GCTTTACACT TTATGCTTCC GGCTCGTATG TTGTGTGGAA	6180
TTGTGAGCGG ATAACAATTT CACACGGGTC ACTTGGCACT GGGCGTCGTT TTACAACGTC	6240
GTGACTGGGA AAACCCTGGC GTTACCCAAG CTTTGTACAT GGAGAAAATA AAGTGAAACA	6300
AAGCACTATT GCACTGGCAC TCTTACCGTT ACCGTTACTG TTTACCCCTG TGACAAAAGC	6360
CGCCGAGGTC CAGCTGCTCG AGTCAGGCCT ATTGTGCCCA GGGGATTGTA CTAGTGGATC	6420
CTAGGCTGAA GCGCATGACC CTGCTAAGGC TGCATTCAAT AGTTTACAGG CAAGTGCTAC	6480
TGAGTACAAT GGCTACGCTT GGGCTATGGT AGTAGTTATA GTTGGTGCTA CCATAGGGAT	6540

TAAATTATTC AAAAAGTTTA CGAGCAAGGC TTCTTAAGCA ATAGCGAAGA GGCCCGCACC	6600
GATCGCCCTT CCCAACAGTT GCGCAGCCTG AATGGCGAAT GGCGCTTTCC CTGGTTTCCG	6660
GCACCAGAAG CCGTGCCGGA AAGCTGGCTG GAGTGGGATC TTCCTGAGGC CGATACGGTC	6720
GTGCTCCCCT CAAACTGGCA GATGCACGGT TACGATGCGC CCATCTACAC CAACGTAACC	6780
TATCCGATTA CCGTCAATCC GCCGTTTGTG CCCACGGAGA ATCCGACGGG TTGTTACTCG	6840
CTCACATTTA ATGTTGATGA AAGCTGGCTA CAGGAAGGCC AGACGCGAAT TATTTTGGAT	6900
GGCGTTCCTA TTGGTTAAAA AATGAGCTGA TTTAACAAAA ATTTAACGCG AATTTTAACA	6960
AAATATTAAC GTTTACAATT TAAATATTTG CTTATACAAT CTTCTGTTTT TTGGGGCTTT	7020
TCTGATTATC AACCGGGGTA CATATGATTG ACATGCTAGT TTTACGATTA CCGTTCATCG	7080
ATTCTCTTGT TTGCTCCAGA CTCTCAGGCA ATGACCTGAT AGCCTTTGTA GATCTCTCAA	7140
AAATAGCTAC CCTCTCCGGC ATTAATTTAT CAGCTAGAAC GGTGAATAT CATATTGATG	7200
GTGATTTGAC TGCTCCGGC CTTTCTCACC CTTTGAATC TTTACCTACA CATTACTCAG	7260
GCATTGCAAT TAAAATATAT GAGGGTTCTA AAAATTTTTA TCCTTGCGTT GAAATAAAGG	7320
CTTCTCCCGC AAAAGTATTA CAGGGTCATA ATGTTTTTGG TACAACCGAT TTAGCTTTAT	7380
GCTCTGAGGC TTTATTGCTT AATTTTGCTA ATTCTTTGCC TTGCCTGTAT GATTATTGG	7440
ACGTT	7445

## (2) INFORMATION FOR SEQ ID NO:2:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 7317 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: both
  - (D) TOPOLOGY: circular

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:2:

AATGCTACTA CTATTAGTAG AATTGATGCC ACCTTTTCAG CTCGGGCCCC AAATGAAAAT	60
ATAGCTAAAC AGGTTATTGA CCATTTGCCA AATGTATCTA ATGGTCAAAC TAAATCTACT	120
CGTTCCGAGA ATTGGGAATC AACTGTTACA TGAATGAAA CTCCAGACA CCGTACTTTA	180
GTTCATATT TAAAACATGT TGAGCTACAG CACCAGATTC AGCAATTAAG CTCTAAGCCA	240
TCCGCAAAAA TGACCTCTTA TCAAAAGGAG CAATTAAAGG TACTGTCTAA TCCTGACCTG	300
TTGGAGTTTG CTTCCGGTCT GGTTCGCTTT GAAGCTCGAA TTAAAACGCG ATATTTGAAG	360
TCTTTCCGGC TTCCTCTTAA TCTTTTGGAT GCAATCCGCT TTGCTTCTGA CTATAATAGT	420
CAGGGTAAAG ACCTGATTTT TGATTTATGG TCATTCTCGT TTTCTGAACT GTTTAAAGCA	480
TTTGAGGGGG ATTCAATGAA TATTTATGAC GATTCCGCAG TATTGGACGC TATCCAGTCT	540
AAACATTTTA CTATTACCCC CTCTGGCAAA ACTTCTTTTG CAAAAGCCTC TCGCTATTTT	600
GGTTTTTATC GTCGTCTGGT AAACGAGCGT TATGATAGTG TTGCTCTTAC TATGCCTCGT	660
AATTCCTTTT GCGGTATGT ATCTGCATTA GTTGAATGTG GTATTCCTAA ATCTCAACTC	720

ATGAATCTTT CTACCTGTAA TAATGTTGTT CCGTTAGTTC GTTTTATTAA CGTAGATTTT	780
TCTTCCCAAC GTCCTGACTG GTATAATGAG CCAGTTCTTA AAATCGCATA AGGTAATTCA	840
CAATGATTAA AGTTGAAATT AAACCATCTC AAGCCCAATT TACTACTCGT TCTGGTGTTT	900
CTCGTCAGGG CAAGCCTTAT TCACTGAATG AGCAGCTTTG TTACGTTGAT TTGGGTAATG	960
AATATCCGGT TCTTGTCAAG ATTACTCTTG ATGAAGGTCA GCCAGCCTAT GCGCCTGGTC	1020
TGTACACCGT TCATCTGTCC TCTTTCAAAG TTGGTCAGTT CCGTTCCCTT ATGATTGAEC	1080
GTCTGGCGCT CGTCCGGCT AAGTAACATG GAGCAGGTG CCGATTTTGA CACAATTTAT	1140
CAGGCGATGA TACAAATCTC CGTTGTACTT TGTTTCGCGC TTGGTATAAT CGCTGGGGGT	1200
CAAAGATGAG TGTTTTAGTG TATTCTTTCG CCTCTTTCGT TTTAGGTTGG TGCCTTCGTA	1260
GTGGCATTAC GTATTTTACC CGTTTAATGG AAACCTCCTC ATGAAAAAGT CTTTAGTCCT	1320
CAAAGCCTCT GTAGCCGTTG CTACCCTCGT TCCGATGCTG TCTTTCGCTG CTGAGGGTGA	1380
CGATCCCGCA AAAGCGGCCT TTAACCTCCT GCAAGCCTCA GCGACCGAAT ATATCGGTGA	1440
TGCGTGGGCG ATGGTTGTTG TCATTGTGCG CGCAACTATC GGTATCAAGC TGTTTAAGAA	1500
ATTCACCTCG AAAGCAAGCT GATAAACCGA TACAATTAAA GGCTCCTTTT GGAGCCTTTT	1560
TTTTTGGAGA TTTTCAACGT GAAAAAATTA TTATTGCAA TTCCTTTAGT TGTTCCCTTC	1620
TATTCTCACT CCGCTGAAAC TGTGAAAGT TGTTTAGCAA AACCCCATAC AGAAAATTCA	1680
TTTACTAACC TCTGGAAAGA CGACAAAAC TTAGATCGTT ACGCTAACTA TGAGGGTTGT	1740
CTGTGGAATG CTACAGGCGT TGAGTTTGT ACTGGTGACC AAACCTCAGT TTACGGTACA	1800
TGGGTTCTTA TTGGGCTTGC TATCCCTGAA AATGAGGGTG GTGGCTCTGA GGGTGGCGGT	1860
TCTGAGGGTG GCGGTTCTGA GGGTGGCGGT ACTAAACCTC CTGAGTACCG TGATACACCT	1920
ATTCGGGGCT ATACTTATAT CAACCCTCTC GACGGCACTT ATCCGCCTGG TACTGAGCAA	1980
AACCCCGCTA ATCCTAATCC TTCTCTTGAG GAGTCTCAGC CTCTTAATAC TTTATGTTT	2040
CAGAATAATA GGTTCGAAA TAGGCAGGGG GCATTAACTG TTTATACGGG CACTGTTACT	2100
CAAGGCACTG ACCCGGTAA AACTTATTAC CAGTACACTC CTGTATCATC AAAAGCCATG	2160
TATGACGCTT ACTGGAACGG TAAATTGAGA GACTGCGCTT TCCATTCTGG CTTTAATGAA	2220
GATCCATTG TTTGTGAATA TCAAGGCCAA TCGTCTGACC TGCCTCAACC TCCTGTCAAT	2280
GCTGGCGGCG GCTCTGGTGG TGGTCTGGT GCGCGCTCTG AGGGTGGTGG CTCTGAGGGT	2340
GGCGGTTCTG AGGGTGGCGG CTCTGAGGGA GCGCGTTCCG GTGGTGGCTC TGGTCCGGT	2400
GATTTTCATT ATGAAAAGAT GGCAAACGCT AATAAGGGGG CTATGACCGA AAATGCCGAT	2460
GAAAACGGCG TACAGTCTGA CGCTAAAGGC AAACCTGATT CTGTGCTAC TGATTACGGT	2520
GCTGCTATCG ATGGTTTCAT TGGTGACGTT TCCGGCCTTG CTAATGGTAA TGGTGCTACT	2580
GGTGATTTTG CTGGCTCTAA TTCCCAAATG GCTCAAGTCG GTGACGGTGA TAATTCACCT	2640
TTAATGAATA ATTTCCGTCA ATATTTACCT TCCCTCCCTC AATCGGTTGA ATGTCCGCT	2700
TTTGTCTTTA GCGCTCGTAA ACCATATCAA TTTTCTATTG ATTGTGACAA AATAAACTTA	2760

TTCCGTGGTG TCTTTGCGTT TCTTTTATAT GTTGCCACCT TTATGTATGT ATTTTCTACG	2820
TTTGCTAACA TACTGCGTAA TAAGGAGTCT TAATCATGCC AGTTCTTTTG GGTATTCCGT	2880
TATTATTGGG TTTCTCGGT TTCCTTCTGG TAACTTTGTT CGGCTATCTG CTTACTTTTC	2940
TTAAAAAGGG CTTCGGTAAG ATAGCTATTG CTAATTCATT GTTCTTGCT CTTATTATTG	3000
GGCTTAACTC AATTCTTG TGTTATCTCT CTGATATTAG CGCTCAATTA CCCTCTGACT	3060
TTGTTCAAGG TGTTCAAGTA ATTCTCCCGT CTAATGCGCT TCCCTGTTTT TATGTTATTC	3120
TCTCTGTAAA GGCTGCTATT TTCATTTTTG ACGTTAAACA AAAAATCGTT TCTTATTG	3180
ATTGGGATAA ATAATATGGC TGTTTATTTT GTAAGTGGCA AATTAGGCTC TGGAAAGACC	3240
CTCGTTAGCG TTGGTAAGAT TCAGGATAAA ATTGTAGCTG GGTGCAAAAT AGCAACTAAT	3300
CTTGATTAA GGCTTCAAAA CCTCCCGCAA GTCGGGAGGT TCGCTAAAAC GCCTCGCGTT	3360
CTTAGAATAC CGGATAAGCC TTCTATATCT GATTGCTTG CTATTGGGCG CGGTAATGAT	3420
TCCTACGATG AAAATAAAAA CGGCTTGCTT GTTCTCGATG AGTGCGGTAC TTGGTTTAA	3480
ACCCGTCTT GGAATGATAA GGAAAGACAG CGGATTATTG ATTGGTTTCT ACATGCTCGT	3540
AAATTAGGAT GGGATATTAT TTTCTTGTT CAGGACTTAT CTATTGTTGA TAAACAGGCG	3600
CGTTCTGCAT TAGCTGAACA TGTTGTTTAT TGTCGTCGTC TGGACAGAAT TACTTTACCT	3660
TTTGTCGGTA CTTTATATTC TCTTATTACT GGCTCGAAAA TGCCTCTGCC TAAATTACAT	3720
GTGGCGTTG TTAAATATGG CGATTCTCAA TTAAGCCCTA CTGTTGAGCG TTGGCTTTAT	3780
ACTGTAAGA ATTTGTATAA CGCATATGAT ACTAAACAGG CTTTTCTAG TAATTATGAT	3840
TCCGGTGTGTT ATTCTTATTT AACGCCTTAT TTATCACACG GTCGGTATTT CAAACCATTA	3900
AATTTAGGTC AGAAGATGAA GCTTACTAAA ATATATTGTA AAAAGTTTTT ACGCGTTCTT	3960
TGTTCTGCGA TTGGATTGTC ATCAGCATTT ACATATAGTT ATATAAGCCA ACCTAAGCCC	4020
GAGGTTAAAA AGGTAGTCTC TCAGACCTAT GATTTTGATA AATCACTAT TGAATCTTCT	4080
CAGCGTCTTA ATCTAAGCTA TCGCTATGTT TTCAAGGATT CTAAGGGAAA ATTAATTAAT	4140
AGCGACGATT TACAGAAGCA AGGTTATTCA CTCACATATA TTGATTATG TACTGTTTCC	4200
ATTAAAAAAG GTAATTCAAA TGAAATTGTT AAATGTAATT AATTTTGTTT TCTTGATGTT	4260
TGTTTCATCA TCTTCTTTTG CTCAGGTAAT TGAAATGAAT AATTCGCCCTC TCGCGGATTT	4320
TGTAAGTTGG TATTCAAAGC AATCAGGCGA ATCCGTTATT GTTCTCCCG ATGTAAGG	4380
TACTGTTACT GTATATTCAT CTGACGTAA ACCTGAAAAT CTACGCAATT TCTTTATTTC	4440
TGTTTACGT GCTAATAATT TTGATATGGT TGTTCAATT CTTCCATAA TTCAGAAGTA	4500
TAATCCAAAC AATCAGGATT ATATTGATGA ATTGCCATCA TCTGATAATC AGGAATATGA	4560
TGATAATTCC GTCCTTCTG GTGGTTTCTT TGTTCCGCAA AATGATAATG TTAATCAAAC	4620
TTTTAAAAAT AATAACGTTT GGGCAAAGGA TTTAATACGA GTTGTCGAAT TGTTTGTA	4680
GTCTAATACT TCTAAATCCT CAAATGTATT ATCTATTGAC GGCTCTAATC TATTAGTTGT	4740
TAGTGCACCT AAAGATATTT TAGATAACCT TCCTCAATTC CTTTCTACTG TTGATTGCGC	4800

AACTGACCAG ATATTGATTG AGGGTTTGAT ATTTGAGGTT CAGCAAGGTG ATGCTTTAGA	4860
TTTTTCATTT GCTGCTGGCT CTCAGCGTGG CACTGTTGCA GGCGGTGTTA ATACTGACCG	4920
CCTCACCTCT GTTTTATCTT CTGCTGGTGG TTCGTTGGGT ATTTTAAATG GCGATGTTTT	4980
AGGGCTATCA GTTCGCGCAT TAAAGACTAA TAGCCATTCA AAAATATTGT CTGTGCCACG	5040
TATTCTTACG CTTTCAGGTC AGAAGGGTTC TATCTCTGTT GGCCAGAATG TCCCTTTTAT	5100
TACTGGTCGT GTGACTGGTG AATCTGCCAA TGTAAATAAT CCATTTTACA CGATTGAGCG	5160
TCAAAATGTA GGTATTTCCA TGAGCGTTTT TCCTGTTGCA ATGGCTGGCG GTAATATTGT	5220
TCTGGATATT ACCAGCAAGG CCGATAGTTT GAGTTCTTCT ACTCAGGCAA GTGATGTTAT	5280
TACTAATCAA AGAAGTATTG CTACAACGGT TAATTTGCGT GATGGACAGA CTCTTTTACT	5340
CGGTGGCCTC ACTGATTATA AAAACACTTC TCAAGATTCT GGCGTACCGT TCCTGTCTAA	5400
AATCCCTTTA ATCGGCCTCC TGTTTAGCTC CCGCTCTGAT TCCAACGAGG AAAGCACGTT	5460
ATACGTGCTC GTCAAAGCAA CCATAGTACG CGCCCTGTAG CGGCGCATTG AGCGCGGGCG	5520
GTGTGGTGGT TACGCGCAGC GTGACCGCTA CACTTGCCAG CGCCCTAGCG CCCGCTCCTT	5580
TCGCTTTCTT CCCTTCCTTT CTCGCCACGT TCGCCGGCTT TCCCCGTCAA GCTCTAAATC	5640
GGGGGCTCCC TTTAGGGTTC CGATTTAGTG GTTTACGGCA CCTCGACCCC AAAAACTTG	5700
ATTTGGGTGA TGGTTCACGT AGTGGGGCCAT CGCCCTGATA GACGGTTTTT CGCCCTTTGA	5760
CGTTGGAGTC CACGTTCTTT AATAGTGGAC TCTTGTCCA AACTGGAACA ACACTCAACC	5820
CTATCTCGGG CTATTCTTTT GATTTATAAG GGATTTTGCC GATTTGGAA CCACCATCAA	5880
ACAGGATTTT CGCCTGCTGG GGCAAACCAG CGTGGACCGC TTGCTGCAAC TCTCTCAGGG	5940
CCAGGCGGTG AAGGGCAATC AGCTGTTGCC CGTCTCGCTG GTGAAAAGAA AAACCACCT	6000
GGCGCCCAAT ACGCAAACCG CCTCTCCCCG CGCGTTGGCC GATTCATTAA TGCAGCTGGC	6060
ACGACAGGTI TCCCGACTGG AAAGCGGGCA GTGAGCGCAA CGCAATTAAT GTGAGTTAGC	6120
TCACTCATTG GGCACCCAG GCTTTACACT TTATGCTTCC GGCTCGTATG TTGTGTGGAA	6180
TTGTGAGCGG ATAACAATTT CACACGCCAA GGAGACAGTC ATAATGAAAT ACCTATTGCC	6240
TACGGCAGCC GCTGGATTGT TATTACTCGC TGCCCAACCA GCCATGGCCG AGCTCGTGAT	6300
GACCCAGACT CCAGATATCC AACAGGAATG AGTGTTAATT CTAGAACCGG TCACTTGCCA	6360
CTGGCCGTG TTTTACAACG TCGTGAAGTG GAAAACCCTG GCGTTACCCA AGCTTAATCG	6420
GCTTGAGAA TTCCCTTTG CCAGCTGGCG TAATAGCGAA GAGGCCCCGA CCGATCGCCC	6480
TTCCCAACAG TTGCGCAGCC TGAATGGCGA ATGGCGCTTT GCCTGGTTTC CGGCACCAGA	6540
AGCGGTGCCG GAAAGCTGGC TGGAGTGCGA TCTTCCTGAG GCCGATACGG TCCTCGTCCC	6600
CTCAAAGTGG CAGATGCACG GTTACGATGC GCCCATCTAC ACCAACGTAA CCTATCCCAT	6660
TACCGTCAAT CCGCCGTTTG TTCCACCGA GAATCCGACG GGTGTTTACT CGCTCACATT	6720
TAATGTTGAT GAAAGCTGGC TACAGGAAGG CCAGACGCGA ATTATTTTG ATGGCGTTCC	6780
TATTGGTTAA AAAATGAGCT GATTTAACA AAATTTAACC CGAATTTTAA CAAAATATTA	6840



ACGTTTACAA TTAAATATT TGCTTATACA ATCTTCCTGT TTTTGGGGCT TTTCTGATTA	6900
TCAACCGGGG TACATATGAT TGACATGCTA GTTTTACGAT TACCGTTCAT CGATTCTCTT	6960
GTTTGCTCCA GACTCTCAGG CAATGACCTG ATAGCCTTTG TAGATCTCTC AAAAATAGCT	7020
ACGCTCTCCG GCATTAAATTT ATCAGCTAGA ACGGTGGAAT ATCATATTGA TGGTGATTTG	7080
ACTGTCTCCG GCCTTTCTCA CCCTTTTGAA TCTTTACCTA CACATTACTC AGGCATTGCA	7140
TTTAAATAT ATGAGGGTTC TAAAAATTTT TATCCTTGCG TTGAAATAAA GGCTTCTCCC	7200
GCAAAAGTAT TACAGGTCA TAATGTTTTT GGTACAACCG ATTTAGCTTT ATGCTCTGAG	7260
GCCTTATTGC TTAATTTTGC TAATTCCTTG CCTTGCTGT ATGATTTATT GGATGTT	7317

## (2) INFORMATION FOR SEQ ID NO:3:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 7729 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: both
  - (D) TOPOLOGY: circular

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AATGCTACTA CTATTAGTAG AATTGATGCC ACCTTTTCAG CTCGGGCCCC AAATGAAAAT	60
ATAGCTAAAC AGGTTATTGA CCATTTCGGA AATGTATCTA ATGGTCAAAC TAAATCTACT	120
CGTTCCGAGA ATTGGGAATC AACTGTTACA TGGAAATGAAA CTTCCAGACA CCGTACTTTA	180
GTTGCATATT TAAAACATGT TGAGCTACAG CACCAGATTC AGCAATTAAG CTCTAAGCCA	240
TCTGCAAAAA TGACCTCTTA TCAAAAGGAG CAATTAAAGG TACTCTCTAA TCCTGACCTG	300
TTGGAGTTTG CTTCGGGTCT GGTTCCGCTTT GAAGCTCGAA TTAAAACGGC ATATTGGAAG	360
TCTTTCCGGC TTCCTCTTAA TCTTTTIGAT GCAATCCGCT TTGCTTCTGA CTATAATAGT	420
CAGGGTAAAG ACCTGATTTT TGATTTATGG TCATTCTCGT TTTCTGAACT GTTTAAAGCA	480
TTTGAGGGGG ATTCAATGAA TATTTATGAC GATTCCGCAG TATTGGACGC TATCCAGTCT	540
AAACATTTTA CTATTACCCC CTCTGGCAA ACTTCTTTTG CAAAAGCCTC TCGCTATTTT	600
GGTTTTTATC GTCGTCTGGT AAACGAGGGT TATGATAGTG TTGCTCTTAC TATGCCTCGT	660
AATTCCTTTT GCGGTTATGT ATCTGCATTA GTTGAATGTG GTATTCCTAA ATCTCAACTG	720
ATGAATCTTT CTACCTGTAA TAATGTTGTT CCGTTAGTTC GTTTTATTAA CGTAGATTTT	780
TCTTCCCAAC GTCCTGACTG GTATAATGAG CCAGTTCTTA AAATCGCATA AGGTAATTCA	840
CAATGATTAA AGTTGAAATT AAACCATCTC AAGCCCAATT TACTACTCGT TCTGGTGTTT	900
GTCGTCAGGG CAAGCCTTAT TCACTGAATG AGCAGCTTTG TTACGTTGAT TTGGGTAATG	960
AATATCCGGT TCTTGTCAAG ATTACTCTTG ATGAAGGTCA GCCAGCCTAT GCGCCTGGTC	1020
TGTACACCGT TCATCTGTCC TCTTCAAAG TTGGTCAGTT CGGTTCCCTT ATGATTGACC	1080
GTCTGCGCCT CGTTCGGGCT AAGTAACATG GAGCAGGTCTG CGGATTTCCA CACAATTTAT	1140
CAGGCGATGA TACAAATCTC CGTTGTACTT TGTTTCGCGC TTGGTATAAT CGCTGGGGGT	1200

CAAAGATGAG TGTITTTAGTG TATTCTTTTCG CCTCTTTTCGT TTTAGGTTGG TGCCTTCGTA	1260
GTGGCATTAC STATTTTACC CGTTTAATGG AAACTTCCTC ATGAAAAAGT CTTTAGTCCT	1320
CAAAGCCTCT GTAGCCGTTG CTACCCTCGT TCCGATGCTG TCTTTCGCTG CTGAGGGTGA	1380
CGATCCCGCA AAAGCGGCCT TTAACCTCCCT GCAAGCCTCA GCGACCGAAT ATATCGGTTA	1440
TGCGTGGGCG ATGGTTGTTG TCATTGTCCG CGCAACTATC GGTATCAAGC TGTTTAAGAA	1500
ATTCACCTCG AAAGCAAGCT GATAAACCGA TACAATTAAA GGCTCCTTTT GGAGCCTTTT	1560
TTTTTGAGA TTTTCAACGT GAAAAATTA TTATTGCAA TTCCTTTAGT TGTTCCTTTC	1620
TATTCTCACT CCGCTGAAAC TGTTGAAAGT TGTTTAGCAA AACCCCATAC AGAAAATTCA	1680
TTTACTAACC TCTGAAAGA CGACAAAAC TTAGATCGTT ACGCTAACTA TGAGGGTTGT	1740
CTGTGGAATG CTACAGGCGT TGTAGTTTGT ACTGGTGACG AAACTCAGTG TTACGGTACA	1800
TGGGTTTCTA TTGGGCTTGC TATCCCTGAA AATGAGGGTG GTGGCTCTGA GGGTGGCGGT	1860
TCTGAGGGTG GCGGTTCTGA GGGTGGCGGT ACTAAACCTC CTGAGTACGG TGATACACCT	1920
ATTCGGGGCT AACTTATAT CAACCCTCTC GACGGCACTT ATCCGCCTGG TACTGAGCAA	1980
AACCCCGCTA ATCCTAATCC TTCTCTTGAG GAGTCTCAGC CTCTTAATAC TTTCAITGTT	2040
CAGAATAATA GGTTCGAAA TAGGCAGGGG GCATTAACTG TTTATACGGG CACTGTTACT	2100
CAAGGCACTG ACCCCGTAA AACTTATTAC CAGTACACTC CTGTATCATC AAAAGCCATG	2160
TATGACGCTT ACTGGAACGG TAAATTCAGA GACTGCGCTT TCCATTCTGG CTTTAATGAA	2220
GATCCATTCTG TTTGTGAATA TCAAGGCCAA TCGTCTGACC TGCCTCAACC TCCTGTCAAT	2280
GCTGGCGGGG GCTCTGGTGG TGGTTCTGGT GGCGGCTCTG AGGGTGGTGG CTCTGAGGGT	2340
GGCGGTTCTG AGGGTGGCGG CTCTGAGGGA GGCGGTTCCG GTGGTGGCTC TGGTTCCGGT	2400
GATTTTGATT ATGAAAAGAT GGCAAACGCT AATAAGGGGG CTATGACCGA AAATGCCGAT	2460
GAAAACGGCG TACAGTCTGA CGCTAAAGGC AAAC TTGATT CTGTGCTAC TGATTACGGT	2520
GCTGCTATCG ATGGTTTCAT TGGTGACGTT TCCGGCCTTG CTAATGGTAA TGGTGCTACT	2580
GGTGATTTTG CTGGCTCTAA TTCCCAAATG GCTCAAGTGG GTGACGGTGA TAATTCACCT	2640
TTAATGAATA ATTCCGCTCA ATATTTACCT TCCCTCCCTC AATCGGTTGA ATGTGCCCCT	2700
TTTGCTTTA GCGCTGGTAA ACCATATGAA TTTTCTATTG ATTGTGACAA AATAAACTTA	2760
TTCCCTGGTG TCTTTGCGTT TCTTTTATAT GTTGCCACCT TTATGTATGT ATTTTCTACG	2820
TTTGCTAACA TACTGCGTAA TAAGGAGTCT TAATCATGCC AGTTCTTTTG GGTATTCCGT	2880
TATTATTGCG TTTCTCGGT TTCCTTCTGG TAACTTTGTT CGGCTATCTG CTTACTTTTC	2940
TTAAAAAGGG CTTCGGTAAG ATAGCTATTG CTATTTCATT GTTTCTTGCT CTTATTATTG	3000
GGCTTAACTC AATTCTTGTG GGTATCTCT CTGATATTAG CGCTCAATTA CCCTCTGACT	3060
TTGTTCAAGG TGTTCACTTA ATTCTCCCGT CTAATGCGCT TCCCTGTTTT TATGTTATTC	3120
TCTCTGTAAA GGCTGCTATT TTCATTTTTG ACGTTAAACA AAAAATCGTT TCTTATTTGG	3180
ATTGGGATAA ATAATATGGC TGTTTATTTT GTAACCTGGCA AATTAGGCTC TGGAAAGACC	3240

CTCGTTAGCG TCGTAAGAT TCAGGATAAA ATTGTAGCTG GGTGCAAAAT AGCAACTAAT	3300
CTTGATTTAA GCCTTCAAAA CCTCCCGCAA GTCCGGAGGT TCGCTAAAAC GCCTCGCGTT	3360
CTTAGAATAC CSGATAAGCC TTCTATATCT GATTTCGCTG CTATTGGGCG CGGTAATGAT	3420
TCCTACGATG AAAATAAAAA CGGCTTGCTT GTTCTCGATG AGTGCGGTAC TTGGTTTAAT	3480
ACCCGTTCTT GGAATGATAA GGAAAGACAG CCGATTATTG ATTGGTTTCT ACATGCTCGT	3540
AAATTAGGAT GGGATATTAT TTTTCTTGTT CAGGACTTAT CTATTGTTGA TAAACAGGCG	3600
CGTTCTGCAT TAGCTGAACA TGTGTTTAT TGTCGTCGTC TGGACAGAAT TACTTTACCT	3660
TTTGTGGTA CTTTATATTC TCTTATTACT GGCTCGAAAA TGCCTCTGCC TAAATTACAT	3720
GTTGGCGTTG TAAATATGG CGATTCTCAA TTAAGCCCTA CTGTTGAGCG TTGGCTTTAT	3780
ACTGGAAGA ATTTGTATAA CGCATATGAT ACTAAACAGG CTTTTCTAG TAATTATGAT	3840
TCCGGTGTTT ATTCTTATTT AACGCCTTAT TTATCACAGG GTCGGTATTT CAAACCATT	3900
AATTTAGGTC AGAAGATGAA GCTTACTAAA ATATATTTGA AAAAGTTTTT ACGCGTCTT	3960
TGTCTTCCGA TTGGATTTGC ATCAGCATT ACATATAGTT ATATAACCCA ACCTAAGCCG	4020
GAGGTTAAAA AGGTAGTCTC TCAGACCTAT GATTTTGATA AATCACTAT TGACTCTTCT	4080
CAGCGTCTTA ATCTAAGCTA TCGCTATGTT TTCAAGGATT CTAAGGGAAA ATTAATTAAT	4140
AGCGACGATT TACAGAAGCA AGGTTATTCA CTCACATATA TTGATTTATG TACTGTTTCC	4200
ATTAAAAAG GTAATTCAAA TGAAATTGTT AAATGTAATT AATTTTGTTT TCTTGATGTT	4260
TGTTTCATCA TCTTCTTTTG CTCAGGTAAT TGAAATGAAT AATTCGCCTC TCGCGGATTT	4320
TGTAACCTGG TATTCAAAGC AATCAGGCCA ATCCGTTATT GTTCTCCCG ATGTAAAAGG	4380
TACTGTTACT GTATATTCAT CTGACGTAA ACCTGAAAAT CTACGCAATT TCTTTATTTT	4440
TGTTTACCT GCTAATAATT TTGATATGGT TGGTTCAATT CCTTCCATAA TTCAGAAGTA	4500
TAATCCAAAC AATCAGGATT ATATTGATGA ATTGCCATCA TCTGATAATC AGGAATATGA	4560
TGATAATTCC GCTCCTTCTG GTGGTTTCTT TGTCCGCAA AATGATAATG TTAATCAAAC	4620
TTTTAAAT AATAACGTTT GGGCAAAGGA TTTAATACGA GTTGTCGAAT TGTTTGTA	4680
GTCTAATACT TCTAAATCCT CAAATGTATT ATCTATTGAC GGCTCTAATC TATTAGTTGT	4740
TAGTGCACCT AAAGATATTT TAGATAACCT TCCTCAATTC CTTTCTACTG TTGATTGCCC	4800
AACTGACCAG ATATTGATTG AGGGTTTGAT ATTTGAGGTT CAGCAAGGTG ATGCTTTAGA	4860
TTTTCAATTT GCTGCTGGCT CTCAGCGTGG CACTGTTGCA GGCGGTGTTA ATACTGACCG	4920
CCTCACCTCT GTTTTATCTT CTGCTGGTGG TTCGTTCCGT ATTTTAAATG GCGATGTTTT	4980
AGGGCTATCA GTTCGGCGAT TAAAGACTAA TAGCCATTCA AAAATATTGT CTGTGCCACG	5040
TATCTTACG CTTTCAGGTC AGAAGGGTTC TATCTCTGTT GGCCAGAATG TCCCTTTTAT	5100
TACTGGTCGT GTGACTGGTG AATCTGCCAA TGTAATAAAT CCATTTTACA CGATTGAGCG	5160
TCAAAATGTA GGTATTTCCA TGAGCGTTTT TCCTGTTGCA ATGGCTGGCG GTAATATTGT	5220
TCTGGATATT ACCAGCAAGG CCGATAGTTT GAGTTCTTCT ACTCAGGCAA GTGATGTTAT	5280

TACTAATCAA AGAAGTATTG CTACAACGGT TAATTTGCGT GATGGACAGA CTCTTTTACT	5340
CGGTGGCCTC ACTGATTATA AAAACACTTC TCAAGATTCT GGGGTACCGT TCCTGTCTAA	5400
AATCCCTTTA ATCGGCCTCC TGTTTAGCTC CCGCTCTGAT TCCAACGAGG AAAGCACGTT	5460
ATACGTGCTC GTCAAAGCAA CCATAGTACG CGCCCTGTAG CGGCGCATT AAGCGGGCGG	5520
GTGTGGTGGT TACGCGCAGC GTGACCGCTA CACTTGCCAG CGCCCTAGCG CCCGCTCCTT	5580
TCGCTTTCTT CCCTTCCTTT CTCGCCACGT TCGCCGGCTT TCGCCGTCAA GCTCTAAATC	5640
GGGGGCTCCC TTTAGGGTTC CGATTTAGTG CTTTACGGCA CCTCGACCCC AAAAACTTG	5700
ATTTGGGTGA TGGTTCACGT AGTGGGCCAT CGCCCTGATA GACGGTTTTT CGCCCTTTGA	5760
CGTTGGAGTC CACGTTCTTT AATAGTGGAC TCTTGTTC AACTGGAACA AACTCAACC	5820
CTATCTCGGG CTATTCTTTT GATTTATAAG GGATTTTGGC GATTTCCGAA CCACCATCAA	5880
ACAGGATTTT CGCCTGCTGG GGCAAACCAG CGTGGACCGC TTGCTGCAAC TCTCTCAGGG	5940
CCAGGCGGTG AAGGGCAATC AGCTGTTGCC CGTCTCGCTG GTGAAAAGAA AAACCACCT	6000
GGCGCCCAAT ACGCAAACCG CCTCTCCCCG CGCGTTGGCC GATTCATTAA TGCAGCTGGC	6060
ACGACAGGTT TCCCGACTGG AAAGCGGGCA GTGAGCGCAA CGCAATTAAT GTGAGTTAGC	6120
TCACTCATT A GGCACCCAG GCTTTACACT TTAIGCTTCC GGCTCGTATG TTGTGTGGAA	6180
TTGTGAGCGG ATAACAATTT CACACGCGTC ACTTGGCACT GGCGTCTGTT TTACAACGTC	6240
GTGACTGGGA AAACCCTGGC GTTACCCAAG CTTTGTACAT GGAGAAAATA AAGTGAACA	6300
AAGCACTATT GCACTGGCAC TCTTACCGTT ACTGTTTACC CCTGTGGCAA AAGCCCAGGT	6360
CCAGCTGCTC GAGTCGGTCT TCCCCCTGGC ACCCTCCTCC AAGAGCACCT CTGGGGGCAC	6420
AGCGGGCCTG GGCTGCCTGG TCAAGACTAA TTCCCCGAAC CGGTGACGGT GTCGTGGAAC	6480
TCAGGCGCCC TGACCAGCGG CGTGCACACC TTCCCGGCTG TCCTACAGTC CTCAGGACTC	6540
TACTCCCTCA GCAGCGTGGT GACCGTGCCC TCCAGCAGCT TGGGCACCCA GACCTACATC	6600
TGCAACGTGA ATCACAAGCC CAGCAACACC AAGGTGGACA AGAAAGCAGA GCCCAAATCT	6660
TGTACTAGTG GATCCTACCC GTACGACGTT CCGGACTACG CTTCTTAGGC TGAAGGCGAT	6720
GACCCTGCTA AGGCTGCATT CAATAGTTTA CAGGCAAGTG CTA CTGAGTA CATTGGCTAC	6780
GCTTGGGCTA TGGTAGTAGT TATAGTTGGT GCTACCATAG GGATTAAATT ATTCAAAAAG	6840
TTTACGAGCA AGGCTTCTTA AGCAATAGCG AAGAGGCCCC CACCGATCGC CCTTCCCAAC	6900
AGTTGCCGAG CCTGAATGGC GAATGGCGCT TTGCCTGGT TCCGGCACCA GAAGCGGTGC	6960
CGGAAAGCTG GCTGGAGTGC GATCTTCCTG AGGCCCATAC GGTGCTCGTC CCCTCAAAC	7020
GGCAGATGCA CGGTTACGAT GCGCCCATCT ACACCAACGT AACCTATCCC ATTACGGTCA	7080
ATCCGCCGTT TGTTCCACG GAGAATCCGA CGGGTTGTTA CTCGCTCACA TTTAATGTTG	7140
ATGAAAGCTG GCTACAGGAA GGCCAGACGC GAATTATTTT TGATGGCGTT CCTATTGGTT	7200
AAAAAATGAG CTGATTTAAC AAAAAATTTA CGCGAATTTT AACAAAATAT TAACGTTTAC	7260
AATTTAAATA TTTGCTTATA CAATCTTCCT GTTTTTGGGG CTTTTCTGAT TATCAACCGG	7320

GGTACATATG ATTGACATGC TAGTTTTACG ATTACCGTTC ATCGATTCTC TTGTTTGCTC	7380
CAGACTCTCA GGCAATGACC TGATAGCCTT TGTAGATCTC TCAAAAATAG CTACCCTCTC	7440
CGGCATTAAT TTATCAGCTA GAACGGTTGA ATATCATATT GATGGTGATT TGA CTGTCTC	7500
CGGCCTTTCT CACCCTTTTG AATCTTTACC TACACATTAC TCAGGCATTG CATTAAAAAT	7560
ATATGAGGGT TCTAAAAATT TTTATCCTTG CGTTGAAATA AAGGCTTCTC CCGCAAAAGT	7620
ATTACAGGGT CATAATGTTT TTGGTACAAC CGATTAGCT TTATGCTCTG AGGCTTTATT	7680
GCTTAATTTT GCTAATTCTT TGCCTTGCCT GTATGATTIA TTGGACGTT	7729

(2) INFORMATION FOR SEQ ID NO:4:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 7557 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: both  
 (D) TOPOLOGY: circular

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:4:

AATGCTACTA CTATTAGTAG AATTGATGCC ACCTTTTCAG CTCGCGCCCC AAATGAAAAAT	60
ATAGCTAAAC AGGTTATTGA CCATTTCGGA AATGTATCTA ATGGTCAAAC TAAATCTACT	120
CGTTCGCAGA ATTGGGAATC AACTGTTACA TGGAATGAAA CTTCCAGACA CCGTACTTTA	180
GTTGCATATT TAAACATGT TGAGCTACAG CACCAGATTC AGCAATTAAG CTCTAAGCCA	240
TCCGCAAAAA TGACCTCTTA TCAAAAGGAG CAATTAAAGG TACTCTCTAA TCCTGACCTG	300
TTGGAGTTTG CTTCCGGTCT GGTTCCGCTT GAAGCTCGAA TTAAACCCCG ATATTTGAAG	360
TCTTTCGGGC TTCCTCTTAA TCTTTTGTAT GCAATCCGCT TTGCTTCTGA CTATAATAGT	420
CAGGGTAAAG ACCTGATTTT TGATTATGG TCATTCTCGT TTTCTGAACT GTTTAAAGCA	480
TTTGAGGGGG ATTCAATGAA TATTTATGAC GATTCCGCAG TATTGGACGC TATCCAGTCT	540
AAACATTTTA CTATTACCCC CTCTGGCAAA ACTTCTTTTG CAAAAGCCTC TCGCTATTTT	600
GGTTTTATC GTCGTCTGGT AAACGAGGGT TATGATAGTG TTGCTCTTAC TATGCCTCGT	660
AATTCCTTTT GCGGTATGT ATCTGCATTA GTTGAATGTG GTATTCCTAA ATCTCAACTG	720
ATGAATCTTT CTACCTGTAA TAAIGTTGTT CCGTTAGTTC GTTTTATTAA CGTAGATTTT	780
TCTTCCCAAC GTCCTGACTG GTATAATGAG CCAGTTCTTA AAATCGCATA AGGTAATTCA	840
CAATGATTAA AGTTGAAATT AAACCATCTC AAGCCCAATT TACTACTCGT TCTGGTGTTT	900
CTCGTCAGGG CAAGCCTTAT TCACTGAATG AGCAGCTTTG TTACGTTGAT TTGGGTAATG	960
AATATCCGGT TCTTGTCAG ATTACTCTTG ATGAAGGTCA GCCAGCCTAT GCGCCTGGTC	1020
TGTACACCGT TCATCTGTCC TCTTTCAAAG TTGGTCAGTT CGGTTCCCTT ATGATTGACC	1080
GTCTGCCGCT CGTCCCGGCT AAGTAACATG GAGCAGGTCT CCGATTTCGA CACAATTTAT	1140
CAGGCGATGA TACAAATCTC CGTTGTACTT TGTTTCGGGC TTGGTATAAT CGCTGGGGGT	1200
CAAAGATGAG TGTTTTAGTG TATTCTTTCT CCTCTTTCGT TTTAGGTTGG TGCCTTCGTA	1260

GTGGCATTAC GTATTTTACC CGTTTAAATGG AAACCTTCCTC ATGAAAAAGT CTTTAGTCCT	1320
CAAAGCCTCT GTAGCCGTTG CTACCCTCGT TCCGATGCTG TCTTTCGCTG CTGAGGGTGA	1380
CGATCCCGCA AAAGCGGCCT TTAACCTCCCT GCAAGCCTCA GCGACCGAAT ATATCGGTGA	1440
TGCGTGGGCG ATGGTTGTTG TCATTGTCGG CGCAACTATC GGTATCAAGC TGTTTAAGAA	1500
ATTCACCTCG AAAGCAAGCT GATAAACCGA TACAATTAAA GGCTCCTTTT GGAGCCTTTT	1560
TTTTTGGA GA TTTCAACGT GAAAAAATTA TTATTGCGAA TTCCTTTAGT TGTTCCCTTC	1620
TATTCTCACT CCGCTGAAAC TGTTGAAAGT TGTTTAGCAA AACCCCATAC AGAAAATTCA	1680
TTTACTAAG TCTGGAAAGA CGACAAAACCT TTAGATCGTT ACGCTAACTA TGAGGGTTGT	1740
CTGTGGAATG CTACAGGCGT TGTAAGTTGT ACTGGTGACG AAACCTCAGT TTACGGTACA	1800
TGGGTTCTTA TTGGGCTTGC TATCCCTGAA AATGAGGGTG GTGGCTCTGA GGGTGGCGGT	1860
TCTGAGGGTG GCGGTTCTGA GGGTGGCGGT ACTAAACCTC CTGAGTACGG TGATACACCT	1920
ATTCCGGGCT ATACTTATAT CAACCCTCTC GACGGCACTT ATCCGCCTGG TACTGAGCAA	1980
AACCCCGCTA ATCCTAATCC TTCTCTTGAG GAGTCTCAGC CTCTTAATAC TTTCATGTTT	2040
CAGAATAATA GGTTCGAAA TAGGCAGGGG GCATTAACTG TTTATACGGG CACTGTTACT	2100
CAAGGCACTG ACCCCGTAA AACTTATTAC CAGTACACTC CTGTATCATC AAAAGCCATG	2160
TATGACGCTT ACTGGAACGG TAAATTCAGA GACTGCGCTT TCCATTCTCG CTTTAATGAA	2220
GATCCATTG TTTGTGAATA TCAAGGCCAA TCGTCTGACC TGCCTCAACC TCCTGTCAAT	2280
GCTGGCGGCG GCTCTGGTGG TGGTCTGGT GGCGGCTCTG AGGGTGGTGG CTCTGAGGGT	2340
GGCGGTTCTG AGGGTGGCGG CTCTGAGGGA GGCGGTTCCG GTGGTGGCTC TGGTTCGGGT	2400
GATTTTGATT ATGAAAAGAT GGCAAACGCT AATAAGGGGG CTATGACCGA AAATGCCGAT	2460
GAAACGCGC TACAGTCTGA CGCTAAAGGC AAACCTTGATT CTGTGCTAC TGATTACGGT	2520
GCTGCTATCG ATGGTTTCAT TGGTCACGTT TCCGGCCTTG CTAATGGTAA TGGTGCTACT	2580
GGTGATTTT CTGGCTCTAA TTCCCAAATG GCTCAAGTCG GTGACGGTGA TAATTCACCT	2640
TTAATGAATA ATTTCGTC AATTTTACCT TCCCTCCCTC AATCGGTTGA ATGTGCGCCT	2700
TTTGTCTTAA GCGCTGGTAA ACCATATGAA TTTTCTATTG ATTGTGACAA AATAAACTTA	2760
TTCCGTGGTG TCTTTGCGTT TCTTTTATAT GTTGCCACCT TTATGTATGT ATTTTCTACG	2820
TTTGCTAACA TACTGCGTAA TAAGGAGTCT TAATCATGCC AGTTCTTTTG GGTATTCCGT	2880
TATTATTGCG TTTCCTCGGT TTCCTTCTGG TAACTTTGTT CCGCTATCTG CTTACTTTTC	2940
TTAAAAAGG CTTCCGTAAG ATAGCTATTG CCTGTTTCTT GCTCTTATTA TTGGGCTTAA	3000
CTCAATTCTT GTGGGTATC TCTCTGATAT TAGCGCTCAA TTACCCTCTG ACTTTGTTC	3060
GGGTGTTTCA TTAATTCTCC CGTCTAATGC GCTTCCCTGT TTTTATGTTA TTCTCTCTGT	3120
AAAGGCTGCT ATTTTCATTT TTGACGTAA ACAAAAAATC GTTCTTATT TGGATTGGGA	3180
TAAATAATAT GGCTGTTTAT TTTGTAAGT GCAAATTAGG CTCTGGAAAG ACGCTCGTTA	3240
GCGTTGGTAA GATTCAGGAT AAAATTGTAG CTGGGTGCAA AATAGCAACT AATCTTGATT	3300

TAAGGCTTCA AAACCTCCCG CAAGTCGGGA GGTTCGCTAA AACGCCTCGC GTTCTTAGAA	3360
TACCGGATAA GCCTTCTATA TCTGATTTGC TTGCTATTGG GCGCGTAAT GATTCCTACG	3420
ATGAAAATAA AAACGGCTTG CTTGTTCTCG ATGAGTGGCG TACTTGGTTT AATACCCGTT	3480
CTTGGAATGA TAAGGAAAGA CAGCCGATTA TTGATTGGTT TCTACATGCT CGTAAATTAG	3540
GATGGGATAT TATTTTCTT GTTCAGGACT TATCTATTGT TGATAAACAG GCGCGTTCTG	3600
CATTAGCTGA ACATGTTGTT TATTGTCGTC GTCTGGACAG AATTACTTTA CCTTTTGTCG	3660
GTACTIONATA TTCTCTTATT ACTGGCTCGA AAATGCCTCT GCCTAAATTA CATGTTGGCG	3720
TTGTTAAATA TGGCGATTCT CAATTAAGCC CTACTGTTGA GCGTTGGCTT TATACTGGTA	3780
AGAATTTGTA TAACGCATAT GATACTAAAC AGGCTTTTTT TAGTAATTAT GATTCGGGTG	3840
TTTATTCTTA TTAAACGGCT TATTATCAC ACGGTCGGTA TTTCAAACCA TTAAATTTAG	3900
GTCAGAAGAT GAAGCTTACT AAAATATATT TGAAAAAGTT TTCACGCGTT CTTTGTCTTG	3960
CGATTGGATT TGCATCAGCA TTTACATATA GTTATATAAC CCAACCTAAG CCGGAGGTTA	4020
AAAAGGTAGT CTCTCAGACC TATGATTTTG ATAAATTCAC TATTGACTCT TCTCAGCGTC	4080
TTAATCTAAG CTATCGCTAT GTTTTCAAGG ATTCTAAGGG AAAATTAATT AATAGCGACG	4140
ATTACAGAA GCAAGGTTAT TCACTCACAT ATATTGATT ATGTACTGTT TCCATTAAAA	4200
AAGGTAATTC AAATGAAATT GTTAAATGTA ATTAATTTTG TTTTCTTGAT GTTTGTTTCA	4260
TCATCTTCTT TTGCTCAGGT AATTGAAATG AATAATTCGC CTCTGCGCGA TTTTGTAAC	4320
TGGTATTCAA AGCAATCAGG CGAATCCGTT ATTGTTTCTC CCGATGTAAA AGGTACTGTT	4380
ACTGTATATT CATCTGACGT TAAACCTGAA AATCTACGCA ATTTCTTTAT TTCTGTTTTA	4440
CGTGCTAATA ATTTTGATAT GGTGTTTCA ATTCCTTCCA TAATTCAGAA GTATAATCCA	4500
AACAATCAGG ATTATATTGA TGAATTGCCA TCATCTGATA ATCAGGAATA TGATGATAAT	4560
TCCGCTCCTT CTGGTGGTTT CTTTGTTCGG CAAAATGATA ATGTTACTCA AACTTTTTAAA	4620
ATTAATAAEC TTCGGGCAAA GGATTTAATA CGAGTTGTGG AATTGTTTGT AAAGTCTAAT	4680
ACTTCTAAAT CCTCAAATGT ATTATCTATT GACGGCTCTA ATCTATTAGT TGTTAGTGCA	4740
CCTAAAGATA TTTAGATAA CCTTCCTCAA TTCCTTTCTA CTGTTGATT GCCAACTGAC	4800
CAGATATTGA TTCAGGGTTT GATATTTGAG GTTCAGCAAG GTGATGCTTT AGATTTTTC	4860
TTTGCTGCTG GGTCTCAGCG TGGCACTGTT GCAGGCGGTG TTAATACTGA CCGCCTCACC	4920
TCTGTTTTAT CTTCTGCTGG TGGTTCGTTT GGTATTTTTA ATGGCGATGT TTTAGGGCTA	4980
TCAGTTCGCG CATTAAAGAC TAATAGCCAT TCAAAAATAT TGTCTGTGCC ACGTATTCTT	5040
ACGCTTTCAG GTCAGAAGGG TTCTATCTCT GTTGGCCAGA ATGTCCCTTT TATTACTGGT	5100
CGTGTGACTG GTGAATCTGC CAATGTAAAT AATCCATTTT AGACGATTGA GCGTCAAAAT	5160
GTAGGTATTT CCATGAGCGT TTTTCTGTT GCAATGGCTG GCGGTAATAT TGTCTGGAT	5220
ATTACCAGCA AGGCCGATAG TTTGAGTTCT TCTACTCAGG CAAGTGATGT TATTACTAAT	5280
CAAAGAAGTA TTGCTACAAC GGTAAATTTT CGTGATGGAC AGACTCTTTT ACTCGGTGGC	5340

CTCACTGATT ATAAAAACAC TTCTCAAGAT TCTGGCGTAC CGTTCCTGTC TAAATCCCT	5400
TTAATCGGCC TCCTGTTTAG CTCCCGCTCT GATTCCAACG AGGAAAGCAC GTTATACGTG	5460
CTCGTCAAAG CAACCATAGT ACGCGCCCTG TAGCGGCGCA TTAAGCGCGG CGGGTGTGGT	5520
GGTTACGGCG ACGGTGACCG CTACACTTGC CAGCGCCCTA GCGCCCGCTC CTTTCGCTTT	5580
CTTCCCTTCC TTTCTCGCCA CGTTCGCGG CTTTCCCGCT CAAGTCTAA ATCGGGGGCT	5640
CCCTTTAGGG TTCCGATTTA GTGCTTTACG GCACCTCGAC CCCAAAAAAC TTGATTGGG	5700
TGATGGTTCA CGTAGTGGG CATCGCCCTG ATAGACGGT TTTCCGCTT TGACGTTGGA	5760
GTCCACGTTT TTTAATAGTG GACTCTTGTT CCAAACCTGGA ACAACACTCA ACCCTATCTC	5820
GGGCTATTCT TTTGATTAT AAGGGATTT GCCGATTTG GAACCACCAT CAAACAGGAT	5880
TTTCGCTGC TGGGGCAAAC CAGCGTGGAC CGCTTGCTGC AACTCTCTCA GGGCCAGGCG	5940
GTGAAGGGCA ATCAGCTGTT GCGCGTCTCG CTGGTGA AAAAACCAC CCTGGCGCCC	6000
AATACGCAA CCGCTCTCC CCGCGCGTTG GCCGATTCAT TAATGCAGCT GGCACGACAG	6060
GTTTCCCGAC TGGAAAGCGG GCAGTGAGCG CAACGCAATT AATGTGAGTT AGCTCACTCA	6120
TTAGGCACCC CAGGCTTTAC ACTTTATGCT TCCGGCTCGT ATGTTGTGTG GAATTGTGAG	6180
CGGATAACAA TTTCACACGC CAAGGAGACA GTCATAATGA AATACCTATT GCCTACGGCA	6240
GCCGCTGGAT TGTTATTACT CGCTGCCCAA CCAGCCATGG CCGAGCTCTT CCGCCCATCT	6300
GATGAGCAGT TGAATCTGG AACTGCCTCT GTTGTGTGCC TGCTGAATAA CTTCTATCCC	6360
AGAGAGGCCA AAGTACAGTG GAAGGTGGAT AACGCCCTCC AATCGGGTAA CTCCCAGGAG	6420
AGTGTACAG AGCAGGACAG CAAGGACAGC ACCTACAGCC TCAGCAGCAC CCTGACGCTG	6480
AGCAAAGCAG ACTACGAGAA ACACAAAGTC TACGCTGCG AAGTCACCCA TCAGGGCCTG	6540
AGCTCGCCCG TCACAAAGAG CTTCAACAGG GGAGAGTGTT CTAGAACCGG TCACCTGGCA	6600
CTGGCCGTCG TTTTACAACG TCGTGA CTGG GAAAACCTG CCGTTACCCA AGCTTAATCG	6660
CCTTGCAGAA TTCCCTTTCC CCAGCTGGCG TAATAGCGAA GAGGCCCGCA CCGATCGCCC	6720
TTCCCAACAG TTGGCGAGCC TGAATGGCGA ATGGCGCTTT GCCTGGTTTC CGGCACCAGA	6780
AGCGGTGCGG GAAAGCTGGC TGGAGTGGCA TCTTCCTGAG GCCGATACGG TCGTCGTCCC	6840
CTCAAACCTGG CAGATGCAGG GTTACGATGC GCCCATCTAC ACCAACGTAA CCTATCCCAT	6900
TACGGTCAAT CCGCCGTTTG TTCCACGGA GAATCCGACG GGTTGTTACT CGCTCAGATT	6960
TAATGTTGAT GAAAGCTGGC TACAGGAAGG CCAGACGCGA ATTATTTTTG ATGGCGTTCC	7020
TATTGGTTAA AAAATGAGCT GATTTAACAA AAATTTAACG CGAATTTTAA CAAAATATTA	7080
ACGTTTACAA TTTAAATATT TGCTTATACA ATCTTCCTGT TTTTGGGGCT TTTCTGATT	7140
TCAACCGGGG TACATATGAT TGACATGCTA GTTTTACGAT TACCGTTCAT CGATTCTCTT	7200
GTTCGCTCCA GACTCTCAGG CAATGACCTG ATAGCCTTTG TAGATCTCTC AAAAATAGCT	7260
ACCTCTCCG GCATTAAATTT ATCAGCTAGA ACGGTTGAAT ATCATATTGA TGGTGAATTG	7320
ACTGTCTCCG GCCTTTCTCA CCCTTTTGAA TCTTTACCTA CACATTACTC AGGCATTGCA	7380



TTTAAATAT ATGAGGGTTC TAAAAATTTT TATCCTTGCG TTGAAATAAA GGCTTCTCCC 7440  
 GCAAAAGTAT TACAGGGTCA TAATGTTTTT GGTACAACCG ATTTAGCTTT ATGCTCTGAG 7500  
 GCTTTATTGC TTAATTTTGC TAATTCTTTG CCTTGCCTGT ATGATTTATT GGATGTT 7557

## (2) INFORMATION FOR SEQ ID NO:5:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8118 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: circular

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AATGCTACTA CTATTAGTAG AATTGATGCC ACCTTTTCAG CTCGGGCCCC AAATGAAAAAT 60  
 ATAGCTAAAC AGGTATTGA CCATTGCGA AATGTATCTA ATGGTCAAAC TAAATCTACT 120  
 CGTTCCGAGA ATTGGGAATC AACTGTTACA TGAATGAAA CTTCCAGACA CCGTACTTTA 180  
 GTTGCAATAT TAAACATGT TGAGCTACAG CACCAGATTC AGCAATTAAG CTCTAAGCCA 240  
 TCTGCAAAAA TGACCTCTTA TCAAAAGGAG CAATTAAAGG TACTCTCTAA TCCTGACCTG 300  
 TTGGAGTTTG CTTCGGTCT GGTTCGCTTT GAAGCTCGAA TTAAACGCG ATATTTGAAG 360  
 TCTTCGGGGC TTCCTCTTAA TCTTTTIGAT GCAATCCGCT TTGCTTCTGA CTATAATAGT 420  
 CAGGGTAAAG ACCTGATTTT TGATTTATGG TCATTCTCGT TTTCTGAACT GTTTAAAGCA 480  
 TTTGAGGGGG ATTCAATGAA TATTTATGAC GATTCCGCAG TATTGGACGC TATCCAGTCT 540  
 AAACATTTTA CTATTACCCC CTCTGGCAA ACTTCTTTTG CAAAAGCCTC TCGTATTTT 600  
 GGTTTTATC GTCGTCTGGT AAACGAGGGT TATGATAGTG TTGCTCTTAC TATGCCTCGT 660  
 AATTCCTTTT GGCCTTATGT ATCTGCATTA GTTCAATGTG GTATTCCTAA ATCTCAACTG 720  
 ATGAATCTTT CTACCTGTAA TAATGTTGTT CCGTTAGTTC GTTTTATTAA CGTAGATTTT 780  
 TCTTCCCAAC<sup>3</sup> GTCCTGACTG GTATAATGAG CCAGTCTTA AAATCGCATA AGGTAATTCA 840  
 CAATGATTAA AGTTCAAAT AAACCATCTC AAGCCCAATT TACTACTCGT TCTGGTGTIT 900  
 CTCGTCAGGG CAAGCCTTAT TCACTGAATG AGCAGCTTTG TIACGTTGAT TTGGGTAATG 960  
 AATATCCGGT TCTTGCAAG ATTACTCTTG ATGAAGGTCA GCCAGCCTAT GCGCCTGGTC 1020  
 TGTACACCGT TCATCTGTCC TCTTCAAAG TTGGTCAGTT CCGTTCCCTT ATGATTGACC 1080  
 GTCTGCGCCT CGTTCCGGCT AAGTAACATG GAGCAGGTCTG CGGATTTCGA CACAATTTAT 1140  
 CAGGCGATGA TACAAATCTC CGTTGTACTT TGTTCGCGC TTGCTATAAT CGCTGGGGGT 1200  
 CAAAGATGAG TGTTTTAGTG TATCTTTTCG CCTCTTTCTG TTTAGGTTGG TGCCTTCGTA 1260  
 GTGGCATTAC GTATTTTACC CGTTTAATGG AAATTCCTC ATGAAAAAGT CTTTAGTCCT 1320  
 CAAAGCCTCT GTAGCCGTTG CTACCCTCGT TCCGATGCTG TCTTTCGCTG CTGAGGGTGA 1380  
 CGATCCCGCA AAAGCGGCCT TTAACCTCCT GCAAGCCTCA GCGACCGAAT ATATCCGTTA 1440  
 TCCGTGGCGG ATGTTGTTG TCATTGTCCG CGCAACTATC GGTATCAAGC TGTTAAGAA 1500

ATTCACCTCG AAAGCAAGCT GATAAACCGA TACAATTAAA GGCTCCTTTT GGAGCCTTTT	1560
TTTTTGGAGA TTTTCAACGT GAAAAAATTA TTATTCGCAA TTCCTTTAGT TGTTCCTTTC	1620
TATTCTCACT CCGCTGAAAC TGTGAAAGT TGTTTAGCAA AACCCCATAC AGAAAATTCA	1680
TTTACTAACG TCTGGAAAGA CGACAAAAC TTAGATCGTT ACGCTAACTA TGAGGGTTGT	1740
CTGTGGAATG CTACAGGCGT TGTAGTTTGT ACTGGTGACG AAACCTCAGTG TTACGGTACA	1800
TGGGTTCCTA TTGGGCTTGC TATCCCTGAA AATGAGGGTG GTGGCTCTGA GGGTGGCGGT	1860
TCTGAGGGTG GCGGTCTGA GGGTGGCGGT ACTAAACCTC CTGAGTACGG TGATACACCT	1920
ATTCCGGGGCT ATACTTATAT CAACCCTCTC GACGGCACTT ATCCGCCTGG TACTGAGCAA	1980
AACCCCGCTA ATCCTAATCC TTCTCTTGAG GAGTCTCAGC CTCTTAATAC TTTCAITTTT	2040
CAGAATAATA GGTTCGAAA TAGGCAGGGG GCATTAAC TG TTTATACGGG CACTGTACT	2100
CAAGGCACTG ACCCGGTAA AACTTATTAC CAGTACACTC CTGTATCATC AAAAGCCATG	2160
TATGACGCTT ACTGGAACGG TAAATTCAGA GACTGCGCTT TCCATTCTGG CTTTAATGAA	2220
GATCCATTCTG TTTGTGAATA TCAAGGCCAA TCGTCTGACC TGCCTCAACC TCCTGTCAAT	2280
GCTGGCGGGG GCTCTGGTGG TGGTTCTGGT GCGGGCTCTG AGGGTGGTGG CTCTGAGGGT	2340
GGCGGTCTCTG AGGGTGGCGG CTCTGAGGGA GCGGGTTCCG GTGGTGGCTC TGGTTCCGGT	2400
GATTTTGATT ATGAAAAGAT GGCAAACGCT AATAAGGGGG CTATGACCGA AAATGCCGAT	2460
GAAAACGCGC TACAGTCTGA CGCTAAAGGC AAACCTTGATT CTGTGCTAC TGATTACGGT	2520
GCTGCTATCG ATGGTTTCAT TGGTGACGTT TCCGGCCTTG CTAATGGTAA TGGTGCTACT	2580
GGTGATTTTG CTGGCTCTAA TTCCCAAATG GCTCAAGTCG GTGACGGTGA TAATTCACCT	2640
TTAATGAATA ATTTCCGTCA ATATTTACCT TCCCTCCCTC AATCGGTTGA ATGTGCCCCCT	2700
TTTGTCTTTA GCGCTGGTAA ACCATATGAA TTTTCTATTG ATTGTGACAA AATAAACTTA	2760
TTCCGTGGTG TCTTTGCGTT TCTTTTATAT GTTGCCACCT TTATGTATGT ATTTTCTACG	2820
TTTGCTAACA TACTGCGTAA TAAGGAGTCT TAATCATGCC AGTTCTTTTG GGTATTCCGT	2880
TATTATTGCG TTTCCTCGGT TTCCTTCTGG TAACTTTGTT CGGCTATCTG CTTACTTTTC	2940
TTAAAAAGGG GTTCGGTAAG ATAGCTATTG GTATTTCAAT GTTCTTGCT CTTATTATTG	3000
GGCTTAACTC AATTCTTGTC GGTATCTCT CTGATATTAG CGCTCAATTA CCCTCTGACT	3060
TTGTTACAGG TGTTACAGTAA ATTCTCCCGT CTAATGCGCT TCCGTGTTTT TATGTTATTC	3120
TCTCTGTAAA GGCTGCTATT TTCATTTTTC ACGTTAAACA AAAAATCGTT TCTTATTTGG	3180
ATTGGGATAA ATAATATGGC TGTTTATTTT GTAACTGGCA AATTAGGCTC TGGAAAGACG	3240
CTCGTTAGCG TTGGTAAGAT TCAGGATAAA ATTGTAGCTG GGTGCAAAAT AGCAACTAAT	3300
CTTGATTITAA GGCTTCAAAA CCTCCCGCAA GTCGGGAGGT TCGCTAAAAAC GCCTCGCGTT	3360
CTTAGAATAC CGGATAAGCC TTCTATATCT GATTTGCTTG CTATTGGGCG CGGTAATGAT	3420
TCCTACGATG AAAATAAAAA CGGCTTGCTT GTTCTCGATG AGTGCGGTAC TTGGTTTAAT	3480
ACCCGTTCTT GGAATGATAA GGAAAGACAG CCGATTATTG ATTGGTTTCT ACATGCTCGT	3540

AAATTAGGAT GGGATATTAT TTTTCTTGTT CAGGACTTAT CTATTGTTGA TAAACAGGCC	3600
CGTTCTGCAT TAGCTGAACA TGTTGTTTAT TGTCGTCGTC TGGACAGAAT TACTTTACCT	3660
TTTGTCCGTA CTTTATATTC TCTTATTACT GGCTCGAAAA TGCCTCTGCC TAAATTACAT	3720
GTTGGCGTTG TTAAATATGG CGATTCTCAA TTAAGCCCTA CTGTTGAGCG TTGGCTTTAT	3780
ACTGGTAAGA ATTTGTATAA CGCATATGAT ACTAAACAGG CTTTTTCTAG TAATTATGAT	3840
TCCGGTGTTT ATTCTTATTT AACGCCTTAT TTATCAGACG GTCGGTATTT CAAACCATTA	3900
AATTTAGGTC AGAAGATGAA GCTTACTAAA ATATATTTGA AAAAGTTTTT ACGCCGTTCTT	3960
TGTCTTGCGA TTGGATTTCG ATCAGCATTT ACATATAGTT ATATAACCCA ACCTAAGCCG	4020
GAGGTTAAAA AGGTAGTCTC TCAGACCTAT GATTTTGATA AATTCAGTAT TGAATCTTCT	4080
CAGCGTCTTA ATCTAAGCTA TCGCTATGTT TTCAAGGATT CTAAGGGAAA ATTAATTAAT	4140
AGCGACGATT TACAGAAGCA AGGTTATTCA CTCACATATA TTGATTTATG TACTGTTTCC	4200
ATTAAAAAAG GTAATTCAAA TGAAATTGTT AAATGTAATT AATTTTGTTT TCTTGATGTT	4260
TGTTTCATCA TCTTCTTTTG CTCAGGTAAT TGAAATGAAT AATTCGCCTC TGCGCGATTT	4320
TGTAACCTGG TATTCAAAGC AATCAGGCGA ATCCGTTATT GTTCTCTCCG ATGTAAAAGG	4380
TACTGTTACT GTATATTCAT CTGACGTTAA ACCTGAAAAT CTACGCAATT TCTTTATTTT	4440
TGTTTTACGT GCTAATAATT TTGATATGGT TGGTTCAATT CCTTCCATAA TTCAGAAGTA	4500
TAATCCAAAC AATCAGGATT ATATTGATGA ATTGCCATCA TCTGATAATC AGGAATATGA	4560
TGATAATTCC GCTCCTTCTG GTGGTTTCTT TGTTCCGCAA AATGATAATG TTAATCAAAC	4620
TTTTAAAAAT AATAACGTTT GGGCAAAGGA TTAAATACGA GTTGTCGAAT TGTTTGTA	4680
GTCTAATACT TCTAAATCCT CAAATGTATT ATCTATTGAC GGCTCTAATC TATTAGTTGT	4740
TAGTGCACCT AAAGATATTT TAGATAACCT TCCTCAATTC GTTCTACTG TTGATTTGCC	4800
AACTGACCAG ATATTGATTG AGGGTTTGAT ATTTGAGGTT CAGCAAGGTG ATGCTTTAGA	4860
TTTTTCATTT GCTGCTGGCT CTCAGCGTGG CACTGTTGCA GGCGGTGTTA ATACTGACCG	4920
CCTCACCTCT GTTTTATCTT CTGCTGGTGG TTCGTTCCGT ATTTTAAATG GCGATGTTTT	4980
AGGGCTATCA GTTCGCGCAT TAAAGACTAA TAGCCATTCA AAAATATTGT CTGTGCCAGG	5040
TATTCTTACG CTTTCAGGTC AGAAGGGTTC TATCTCTGTT GGCCAGAATG TCCCTTTTAT	5100
TACTGGTCGT GTGACTGGTG AATCTGCCAA TGTAATAAAT CCATTTTACA CGATTGAGCG	5160
TCAAAATGTA GGTATTTCCA TGAGCGTTTT TCCTGTTGCA ATGGCTGGCG GTAATATTGT	5220
TCTGGATATT ACCAGCAAGG CCGATAGTTT GAGTTCTTCT ACTCAGGCAA GTGATGTTAT	5280
TACTAATCAA AGAAGTATTG CTACAACGGT TAATTGCGT GATGGACAGA CTCTTTTACT	5340
CGGTGGCCTC ACTGATTATA AAAACACTTC TCAAGATTCT GGCGTACCGT TCCTGTCTAA	5400
AATCCCTTTA ATCGGCCTCC TGTTTAGCTC CCGCTCTGAT TCCAACGAGG AAAGCACGTT	5460
ATACGTGCTC GTCAAAGCAA CCATAGTACG CGCCGTGTAG CGCGGCATTA AGCGCGGCGG	5520
GTGTGGTGGT TACGCGCAGC GTGACCGCTA CACTTGCCAG CGCCCTAGCG CCCGCTCCTT	5580

TCGCTTTCTT CCGTTCTTT CTGCCCACGT TCGCCGGCTT TCCCCGTCAA GCTCTAAATC	5640
GGGGGCTCCC TTTAGGGTTC CGATTTAGTG CTTTACGGCA CCTCGACCCC AAAAACTTG	5700
ATTTGGGTGA TGGTTCACGT AGTGGGCCAT CGCCCTGATA GACGGTTTTT CGCCCTTTGA	5760
CGTTGGAGTC CACGTTCTTT AATAGTGGAC TCTTGTCCA AACTGGAACA AACTCAACC	5820
CTATCTCGGG CTATTCTTTT GATTTATAAG GGATTTTGCC GATTTGCGAA CCACCATCAA	5880
ACAGGATTTT CGCCTGCTGG GGCAAACCAG CGTGGACCGC TTGCTGCAAC TCTCTCAGGG	5940
CCAGGCGGTG AAGGGCAATC AGCTGTTGCC CGTCTCGCTG GTGAAAAGAA AAACCACCCT	6000
GGCGCCCAAT ACGCAAACCG CCTCTCCCCG CGCGTTGGCC GATTCATTAA TGCAGCTGGC	6060
ACGACAGGTT TCCCGACTGG AAAGCGGGCA GTGAGCGCAA CGCAATTAAT GTGAGTTAGC	6120
TCACTCATTG GGCACCCAG GCTTTACACT TTATGCTTCC GGCTCGTATG TTGTGTGGAA	6180
TTGTGAGCGG ATAACAATTT CACACGCCAA GGAGACAGTC ATAATGAAAT ACCTATTGCC	6240
TACCGCAGCC GCTGGATTGT TATTACTCGC TGCCCAACCA GCCATGGCCG AGCTCTTCCC	6300
GCCATCTGAT GAGCAGTTGA AATCTGGAAC TGCTCTGTT GTGTGCCTGC TGAATAACTT	6360
CTATCCCAGA GAGGCCAAAG TACAGTGGAA GGTGGATAAC GCCCTCCAAT CGGGTAACTC	6420
CCAGGAGAGT GTCACAGAGC AGGACAGCAA GGACAGCACC TACAGCCTCA GCAGCACCCT	6480
GACGCTGAGC AAAGCAGACT ACGAGAAACA CAAAGTCTAC GCCTGCGAAG TCACCCATCA	6540
GGGCTGAGC TCGCCCGTCA CAAAGAGCTT CAACAGGGGA GAGTGTCTA GAACGCGTCA	6600
CTTGGCACTG GCCGTCGTTT TACAACGTCG TGAAGTGGAA AACCCTGGCG TTACCCAAGC	6660
TTTGTACATG GAGAAAATAA AGTGAACAA AGCACTATTG CACTGGCACT CTTACCGTTA	6720
CTGTTTACCC CTGTGGCAA AGCCGCTCC ACCAAGGGCC CATCGGTCTT CCCCCTGGCA	6780
CCCTCTCCA AGAGCACCTC TGGGGGCACA GCGGCCCTGG GCTGCCTGGT CAAGACTAAT	6840
TCCCCGAACC GGTGACGGTG TCGTGGAAT CAGGCGCCCT GACCAGCGGC GTGCACACCT	6900
TCCCGGCTGT CCTACAGTCC TCAGGACTCT ACTCCCTCAG CAGCGTGGTG ACCGTGCCCT	6960
CCAGCAGCTT GGGCACCCAG ACCTACATCT GCAACGTGAA TCACAAGCCC AGCAACACCA	7020
AGGTGGACAA GAAAGCAGAG CCCAAATCTT GTACTAGTGG ATCCTACCCG TACGACGTTT	7080
CGGACTACGC TTCTTAGGCT GAAGGCGATG ACCCTGCTAA GGCTGCAATC AATAGTTTAC	7140
AGGCAAGTGC TACTGAGTAC ATTGGCTACG CTTGGGCTAT GGTAGTAGTT ATAGTTGGTG	7200
CTACCATAGG GATTAAATTA TTCAAAAAGT TTACGAGCAA GGCTTCTTAA GCAATAGCGA	7260
AGAGGCCCCG ACCGATCGCC GTTCCCAACA GTTGGCAGC CTGAATGGCG AATGGCGCTT	7320
TGCCTGGTTT CCGGCACCAG AAGCGGTGCC GGAAAGCTGG CTGGAGTGGC ATCTTCTGA	7380
GGCCGATACG GTCGTCGTCC CCTCAAACTG GCAGATGCAC GGTTACGATG CGCCCATCTA	7440
CACCAACGTA ACCTATCCCA TTACGGTCAA TCCGCCGTTT GTTCCCACGG AGAATCCGAC	7500
GGGTTGTTAC TCGCTCACAT TTAATGTTGA TGAAAGCTGG CTACAGGAAG GCCAGACGGC	7560
AATTATTTTT GATGGCGTTC CTATTGGTTA AAAAATGAGC TGATTTAACA AAAATTTAAC	7620

GCGAATTTTA AAAAAATATT AACGTTTACA ATTTAAATAT TTGCTTATAC AATCTTCCTG	7680
TTTTTGGGGC TTTTCTGATT ATCAACCGGG GTACATATGA TTGACATGCT AGTTTTACGA	7740
TTACCGTTCA TCGATTCTCT TGTITGCTCC AGACTCTCAG GCAATGACCT GATAGCCTTT	7800
GTAGATCTCT CAAAAATAGC TACCCTCTCC GGCATTAATT TATCAGCTAG AACGGTTGAA	7860
TATCATATTG ATGGTGATTT GACTGTCTCC GGCCTTTCTC ACCCTTTTGA ATCTTTACCT	7920
ACACATTACT CAGGCATTGC ATTTAAAATA TATGAGGGTT CTAAAAATTT TTATCCTTGG	7980
GTTGAAATAA AGGCTTCTCC CGCAAAAGTA TTACAGGGTC ATAATGTTTT TGGTACAACC	8040
GATTTAGCTT TATGCTCTGA GGCTTTATTG CTTAATTTTG CTAATTCTTT GCCTTGCCTG	8100
TATGATTTAT TGGACGTT	8118

## (2) INFORMATION FOR SEQ ID NO:6:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 22 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (ix) FEATURE:

- (A) NAME/KEY: misc\_difference
- (B) LOCATION: replace(5, "")
- (D) OTHER INFORMATION: /note= "S REPRESENTS EQUAL MIXTURE OF G AND C"

## (ix) FEATURE:

- (A) NAME/KEY: misc\_difference
- (B) LOCATION: replace(6, "")
- (D) OTHER INFORMATION: /note= "M REPRESENTS EQUAL MIXTURE OF A AND C"

## (ix) FEATURE:

- (A) NAME/KEY: misc\_difference
- (B) LOCATION: replace(8, "")
- (D) OTHER INFORMATION: /note= "R REPRESENTS EQUAL MIXTURE OF A AND G"

## (ix) FEATURE:

- (A) NAME/KEY: misc\_difference
- (B) LOCATION: replace(11, "")
- (D) OTHER INFORMATION: /note= "K REPRESENTS EQUAL MIXTURE OF G AND T"

## (ix) FEATURE:

- (A) NAME/KEY: misc\_difference
- (B) LOCATION: replace(20, "")
- (D) OTHER INFORMATION: /note= "W REPRESENTS EQUAL MIXTURE OF A AND T"

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:6:

AGGTSMARCT KCTCCAGTCW GG

## (2) INFORMATION FOR SEQ ID NO:7:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 22 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:7:

AGGTCCAGCT GCTCGAGTCT GG

22

## (2) INFORMATION FOR SEQ ID NO:8:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 22 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:8:

AGGTCCAGCT GCTCGAGTCA GG

22

## (2) INFORMATION FOR SEQ ID NO:9:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 22 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:9:

AGGTCCAGCT TCTCGAGTCT GG

22

## (2) INFORMATION FOR SEQ ID NO:10:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 22 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:10:

AGGTCCAGCT TCTCGAGTCA GG

22

## (2) INFORMATION FOR SEQ ID NO:11:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 22 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:11:  
AGGTCCAACT GCTCGAGTCT GG

22

(2) INFORMATION FOR SEQ ID NO:12:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 22 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:12:  
AGGTCCAACT GCTCGAGTCA GG

22

(2) INFORMATION FOR SEQ ID NO:13:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 22 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:13:  
AGGTCCAACT TCTCGAGTCT GG

22

(2) INFORMATION FOR SEQ ID NO:14:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 22 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:14:  
AGGTCCAACT TCTCGAGTCA GG

22

(2) INFORMATION FOR SEQ ID NO:15:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 22 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(1x) FEATURE:

- (A) NAME/KEY: misc\_difference
- (B) LOCATION: replace(5..6, "")
- (D) OTHER INFORMATION: /note= "N-INOSINE"

(1x) FEATURE:

- (A) NAME/KEY: misc\_difference
- (B) LOCATION: replace(8, "")
- (D) OTHER INFORMATION: /note= "N-INOSINE"

- (1x) FEATURE:  
 (A) NAME/KEY: misc\_difference  
 (B) LOCATION: replace(11, "")  
 (D) OTHER INFORMATION: /note- "N-INOSINE"

- (1x) FEATURE:  
 (A) NAME/KEY: misc\_difference  
 (B) LOCATION: replace(20, "")  
 (D) OTHER INFORMATION: /note- "W REPRESENTS EQUAL MIXTURE OF A AND T"

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:15:

AGGTNNANCT NCTCGAGTCW GG

22

(2) INFORMATION FOR SEQ ID NO:16:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 38 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CTATTAATA GTAACGGTAA CAGTGGTGCC TTGCCCCA

38

(2) INFORMATION FOR SEQ ID NO:17:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 30 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:17:

AGGCTTACTA GTACAATCCC TGGGCACAAT

30

(2) INFORMATION FOR SEQ ID NO:18:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 32 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CCAGTTCGGA GCTCGTTGTG ACTCAGGAAT CT

32

(2) INFORMATION FOR SEQ ID NO:19:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 32 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear



(x1) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CCAGTTCGGA GCTCGTGTTG ACCGAGCCGC CC

32

(2) INFORMATION FOR SEQ ID NO:20:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 32 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CCAGTTCGGA GCTCGTGCTC ACCCAGTCTC CA

32

(2) INFORMATION FOR SEQ ID NO:21:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 32 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CCAGTTCGGA GCTCCAGATG ACCCAGTCTC CA

32

(2) INFORMATION FOR SEQ ID NO:22:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 32 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:22:

CCAGATGTGA GCTCGTGATG ACCCAGACTC CA

32

(2) INFORMATION FOR SEQ ID NO:23:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 32 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CCAGATGTGA GCTCGTCATG ACCCAGTCTC CA

32

(2) INFORMATION FOR SEQ ID NO:24:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 32 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:24:

CCAGTTCCGA GCTCGTGATG ACACAGTCTC CA

32

(2) INFORMATION FOR SEQ ID NO:25:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 32 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GCAGCATTCT AGAGTTTCAG CTCCAGCTTG CC

32

(2) INFORMATION FOR SEQ ID NO:26:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 34 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GCGCCGTCTA GAATTAACAC TCATTCTGT TGAA

34

(2) INFORMATION FOR SEQ ID NO:27:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 37 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi)<sup>3</sup> SEQUENCE DESCRIPTION: SEQ ID NO:27:

GATCCTAGGC TGAAGGCGAT GACCCTGCTA AGGCTGC

37

(2) INFORMATION FOR SEQ ID NO:28:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 35 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:28:

ATTCAATAGT TTACAGGCAA GTGCTACTCA GTACA

35

## (2) INFORMATION FOR SEQ ID NO:29:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 35 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:29:

TTGGCTACGC TTGGGCTATG GTAGTAGTTA TAGTT

35

## (2) INFORMATION FOR SEQ ID NO:30:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 35 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GGTGCTACCA TAGGGATTAA ATTATTCAAA AAGTT

35

## (2) INFORMATION FOR SEQ ID NO:31:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:31:

TACGAGCAAG GCTTCTTA

18

## (2) INFORMATION FOR SEQ ID NO:32:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 39 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:32:

AGCTTAAGAA GCCTTGCTCG TAAACTTTTT GAATAATTT

39

## (2) INFORMATION FOR SEQ ID NO:33:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 36 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:33:

AATCCCTATG GTAGCACCAA CTATAACTAC TACCAT

36

(2) INFORMATION FOR SEQ ID NO:34:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 35 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:34:

AGCCCAAGCG TAGCCAATGT ACTCAGTAGC ACTTC

35

(2) INFORMATION FOR SEQ ID NO:35:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 34 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:35:

CCTGTAAACT ATTGAATGCA GCCTTAGCAG GGTC

34

(2) INFORMATION FOR SEQ ID NO:36:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 16 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:36:

ATCGCCTTGA GCCTAG

16

(2) INFORMATION FOR SEQ ID NO:37:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:37:

CATTTTGGCA GATGGCTTAG A

21

(2) INFORMATION FOR SEQ ID NO:38:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:38:

TAGCATTAAAC GTCCAATA

18

(2) INFORMATION FOR SEQ ID NO:39:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 26 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:39:

ATATATTTTGA GTAAGCTTCA TCTTCT

26

(2) INFORMATION FOR SEQ ID NO:40:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 23 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:40:

GACAAAGAAC GCGTGAAAAC TTT

23

(2) INFORMATION FOR SEQ ID NO:41:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 35 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:41:

GCGGGCCTCT TCGCTATTGC TTAAGAAGCC TTGCT

35

(2) INFORMATION FOR SEQ ID NO:42:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 43 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:42:

AAACGACGGC CAGTGCCAAG TCACGGGTGT GAAATTGTTA TCC

43

## (2) INFORMATION FOR SEQ ID NO:43:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 43 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:43:

GGCGAAAGGG AATTCTGCAA GGCGATTAAAG CTGGGTAAC GCC

43

## (2) INFORMATION FOR SEQ ID NO:44:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 36 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:44:

GGCGTTACCC AAGCTTTGTA CATGGAGAAA ATAAAG

36

## (2) INFORMATION FOR SEQ ID NO:45:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 42 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:45:

TGAAACAAAG CACTATTGCA CTGGCACTCT TACCGTTACC GT

42

## (2) INFORMATION FOR SEQ ID NO:46:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 42 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:46:

TACTGTTTAC CCCTGTGACA AAAGCCGCCC AGGTCCAGCT GC

42

## (2) INFORMATION FOR SEQ ID NO:47:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 44 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:47:

TCGAGTCAGG CCTATTGTGC CCAGGGATTG TACTAGTGA TCCG

44

(2) INFORMATION FOR SEQ ID NO:48:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 38 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:48:

TGGCGAAAGG GAATTCGGAT CCACTAGTAC AATCCCTG

38

(2) INFORMATION FOR SEQ ID NO:49:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 42 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:49:

GGCACAATAG GCCTGACTCG AGCAGCTGGA CCAGGGCGGC TT

42

(2) INFORMATION FOR SEQ ID NO:50:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 42 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:50:

TTGTCACAGG GGTAAACAGT AACGGTAACG GTAAGTGTGC CA

42

(2) INFORMATION FOR SEQ ID NO:51:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 42 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:51:

GTGCAATAGT GCTTTGTTTC ACTTTATTTT CTCCATGTAC AA

42

(2) INFORMATION FOR SEQ ID NO:52:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:52:  
TAACGGTAAG AGTGCCAGTG C

21

(2) INFORMATION FOR SEQ ID NO:53:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 32 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:53:  
CACCTTCATG AATTCGGCAA GGAGACAGTC AT

32

(2) INFORMATION FOR SEQ ID NO:54:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 22 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:54:  
AATTCGCCAA GGAGACAGTC AT

22

(2) INFORMATION FOR SEQ ID NO:55:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 39 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:55:  
AATGAAATAC CTATTGCCTA CGGCAGCCGC TGGATTGTT

39

(2) INFORMATION FOR SEQ ID NO:56:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 39 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:56:  
ATTACTCGCT GCCCAACCAG CCATGCCCGA GCTCGTGAT

39



(2) INFORMATION FOR SEQ ID NO:57:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 39 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:57:

GACCCAGACT CCAGATATCC AACAGGAATC AGTGTTAAT

39

(2) INFORMATION FOR SEQ ID NO:58:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 13 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:58:

TCTAGAACGC GTC

13

(2) INFORMATION FOR SEQ ID NO:59:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 45 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:59:

TTCAGGTTGA AGCTTACGCG TTCTAGAATT AACACTCATT CCTGT

45

(2) INFORMATION FOR SEQ ID NO:60:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 39 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:60:

TGGATAICTG GAGTCTGGGT CATCAGGAGC TCGGCCATG

39

(2) INFORMATION FOR SEQ ID NO:61:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 39 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:61:

GCTGGTTGGG CAGCGAGTAA TAACAATCCA GCGGCTGCC

39

(2) INFORMATION FOR SEQ ID NO:62:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 37 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:62:

GTAGGCAATA GGTATTTTCAT TATGACTGTC CTTGGCG

37

(2) INFORMATION FOR SEQ ID NO:63:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 30 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:63:

TGACTGTCTC CTTGGCGTGT GAAATGTGA

30

(2) INFORMATION FOR SEQ ID NO:64:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 36 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:64:

TAACACTCAT TCCGGATGGA ATTCTGGAGT CTGGGT

36

(2) INFORMATION FOR SEQ ID NO:65:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:65:

GCCAGTGCCA AGTGACGCGT TCTA

24

(2) INFORMATION FOR SEQ ID NO:66:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 26 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:66:

ATATATTTTGA GTAAGCTTCA TCTTCT

(2) INFORMATION FOR SEQ ID NO:67:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 23 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

26

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:67:

GACAAAGAAC GCGTGAAAAC TTT

(2) INFORMATION FOR SEQ ID NO:68:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 76 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

23

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:68:

CTGAACCTGT CTGGGACCAC AGTTGATGCT ATAGGATCAG ATCTAGAATT CATTAGAGA  
CTGGCCTGGC TTCTGC

60

(2) INFORMATION FOR SEQ ID NO:69:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 80 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

76

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:69:

TCGACCGTTG GTAGGAATAA TGCAATTAAAT GGAGTAGCTC TAAATTCAGA ATTCATCTAC  
ACCCAGTGCA TCCAGTAGCT

60

(2) INFORMATION FOR SEQ ID NO:70:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 27 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

80

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:70:

GGTAAACAGT AACGGTAAGA GTGCCAG

27

## (2) INFORMATION FOR SEQ ID NO:71:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 54 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:71:

CGCCTTCAGC CTAAGAAGCG TAGTCCGGAA CGTCGTACGG GTAGGATCCA CTAG

54

## (2) INFORMATION FOR SEQ ID NO:72:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 41 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:72:

CACCGGTTCC GGAATTAGT CTGACCAGG CAGCCCAGGG C

41

## (2) INFORMATION FOR SEQ ID NO:73:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 51 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:73:

ATTCCACACA TTATACGAGC CGGAAGCATA AAGTGTCAG CCTGGGGTGC C

51

## (2) INFORMATION FOR SEQ ID NO:74:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 42 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:74:

CTGCTCATCA GATGGCGGGA AGAGCTCGGC CATGGCTGGT TG

42

## (2) INFORMATION FOR SEQ ID NO:75:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 42 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

78

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:75:

GAACAGAGTG ACCGAGGGGG CGAGCTCGGC CATGGCTGGT TG

42